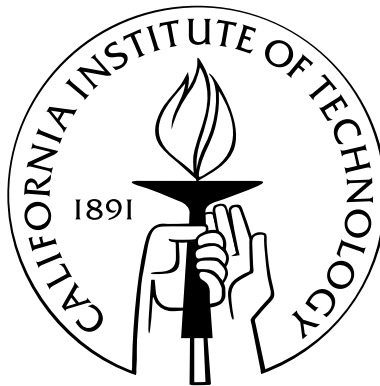


Dynamics of Spinning Compact Binaries in General Relativity

Thesis by
Michael David Hartl

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy



California Institute of Technology
Pasadena, California

2003
(Defended May 22, 2003)

© 2003

Michael David Hartl

All Rights Reserved

Acknowledgements

Earning a Ph.D. in physics from Caltech represents the fulfillment of a lifelong dream, and there are many who have contributed to its realization. First are my mother and father, whose encouragement, love, and support I could take for granted—if only all children were so lucky. It was they, more than anyone else, who were responsible for my early science education, especially through their indulgence of my voracious appetite for popular science (although my wonderful step-mother deserves considerable blame for this indulgence as well). The vacuum created by uninspiring science teachers was filled by books, from authors such as Timothy Ferris, John Casti, Douglas Hofstadter, and Stephen Hawking. My deepest gratitude in this vein is to Carl Sagan, whose *Cosmos* book and miniseries were a revelation. *Cosmos* sparked, among other things, an interest in amateur astronomy, but pursuing this interest might never have happened without the intervention of Lang Dana, whose generous and utterly unexpected gift of a first-rate telescope came—as he well knew—at a critical juncture in my life; for this I shall be eternally grateful.

As an undergraduate at Harvard, it was my great privilege to learn physics from David Bear, who, though then a lowly graduate student, was the best teacher I could have asked for. He was the role model for my own teaching at Caltech, and I hope I have lived up to the lofty standard he set. My other best teachers were my friends, especially Ranjit Aiyagari, Long Nguyen, and Sumit Daftuar (who, among other things, was an ebullient and insightful physics study partner). Perhaps my most valued friend at Harvard was someone I actually met long before—my sister Kara, the best sister a brother could have, and soon to become a *real* doctor.

My journey to Caltech started in the Solar and Stellar X-Ray group at the Harvard-Smithsonian Center for Astrophysics, where I had the good fortune of working with Leon Golub and Ed DeLuca (as well as fellow student and valued friend Meredith Wills). After the end of my sophomore year, I hoped to spend the summer doing research in my home state of California, so the question was obvious: I asked Leon, “Do you know anyone at Caltech?” The answer was yes, and Leon’s old friend Hal Zirin managed to get me a Summer Undergraduate Research Fellowship, even though the application deadline had already long passed. Though I lived far from campus that summer (in the idyllic mountain resort town of Big Bear), then-recent Caltech graduate Glenn Eychaner thoroughly indoctrinated me into the Way of Tech. When I returned to the SURF program for a

second summer, this time living on campus, I already felt like an honorary Techer. I owe a debt of gratitude to Dale Gary for making that second summer possible, and for his patient guidance of the SURF project that led to my first scientific publication.

Much of this thesis draws on the results and techniques of chaos theory, and it was my great privilege to learn this subject from one of its masters, Jim Yorke. I would like to thank the University of Maryland for the Distinguished Research Fellowship program that made that study possible. To Jim Yorke my gratitude is profound, not only for his generosity and wisdom, but also for writing me a strong letter of recommendation to Caltech—even though it meant helping me move to a rival institution.

At Caltech, my friends have been a source of encouragement, support, inspiration, and occasionally beer. Many thanks go especially to Kashif Alvi, Sumit Daftuar, Samantha Edgington, Robert Forster, Scott Hughes, Keith Matthews, Shanti Rao, Kevin Scaldeferri, Luke Sollitt, Chip Sumner, and Shimpei Yamashita.

Of my teachers, I especially want to thank Kip Thorne. I will always feel lucky that I had the opportunity to learn general relativity from the man who put the T in MTW—and refuses to be called anything other than Kip. Much thanks also to Alessandra Buonanno, who supplied the copious notes and references that made Chapter 4 possible. I look forward to collaborating with her on an expanded form of that project. My deepest appreciation goes to my advisor, Sterl Phinney, who was willing (and able!) to indulge my desire to find a project mixing chaos theory and general relativity. It was he who waved the magic wand, seemingly possessed by so many at Caltech, that created a connection between that project and gravitational waves. I am particularly grateful for Sterl’s careful reading of—and resulting insightful comments on—the papers composing this thesis, which always led to major improvements in both content and presentation. Though technically single-author works, the chapters in this thesis owe much to his guidance and advice.

This thesis was supported financially by Caltech teaching assistantships (including the Richard Feynman Book Fund Teaching Assistantship) and by NASA grant NAG5-10707.

Abstract

This thesis investigates the dynamics of binary systems composed of spinning compact objects (such as white dwarfs, neutron stars, and black holes) in the context of general relativity. In particular, we use the method of Lyapunov exponents to determine whether such systems are chaotic. Compact binaries are promising sources of gravitational radiation for both ground- and space-based gravitational-wave detectors, and radiation from chaotic orbits would be difficult to detect and analyze. For chaotic orbits, the number of waveform templates needed to match a given gravitational-wave signal would grow exponentially with increasing detection sensitivity, rendering the preferred *matched filter* detection method computationally impractical. It is therefore urgent to understand whether the binary dynamics can be chaotic, and, if so, how prevalent this chaos is.

We first consider the dynamics of a spinning compact object orbiting a much more massive rotating black hole, as modeled by the Papapetrou equations in Kerr spacetime. We find that many initial conditions lead to positive Lyapunov exponents, indicating chaotic dynamics. The Lyapunov exponents come in positive/negative pairs, a characteristic of Hamiltonian dynamical systems. Despite the formal existence of chaotic solutions, we find that chaos occurs only for physically unrealistic values of the small body's spin. As a result, chaos will not affect theoretical templates in the extreme mass-ratio limit for which the Papapetrou equations are valid. Chaos will therefore not affect the ability of space-based gravitational-wave detectors (such as LISA, the Laser Interferometer Space Antenna) to perform precision tests of general relativity using extreme mass-ratio inspirals.

We next consider the dynamics of spinning black-hole binaries, as modeled by the post-Newtonian (PN) equations, which are valid for orbital velocities much smaller than the speed of light. We study thoroughly the special case of quasi-circular orbits with comparable mass ratios, which are particularly relevant from the perspective of gravitational wave generation for LIGO (the Laser Interferometer Gravitational-wave Observatory) and other ground-based interferometers. In this case, unlike the extreme mass-ratio case, we find chaotic solutions for physically realistic values of the spin. On the other hand, our survey shows that chaos occurs in a negligible fraction of possible configurations, and only for such small radii that the PN approximation is likely to be invalid. As a result, at least in the case of comparable mass black-hole binaries, theoretical templates will not

be significantly affected by chaos.

In a final, self-contained chapter, we discuss various methods for the calculation of Lyapunov exponents in systems of ordinary differential equations. We introduce several new techniques applicable to constrained dynamical systems, developed in the course of studying the dynamics of spinning compact binaries.

Considering the Papapetrou and post-Newtonian systems together, our most important general conclusion is that we find no chaos in any relativistic binary system for orbits that clearly satisfy the approximations required for the equations of motion to be physically valid.

Contents

Acknowledgements	iii
Abstract	v
1 Introduction	1
1.1 The Papapetrou and post-Newtonian equations	4
1.1.1 Papapetrou (or Dixon)	4
1.1.2 Post-Newtonian (with spin)	6
1.1.3 Normalized units and μM	7
1.1.4 Normalization constraints	8
1.1.5 Spin supplementary conditions and the center of mass	9
1.1.6 Degrees of freedom	11
1.2 Lyapunov exponents and chaos	12
1.2.1 Deviation vector method	12
1.2.2 Jacobian method	13
1.2.3 The origin of exponential separation	13
1.2.4 Constrained systems	14
1.2.5 The meaning of zero	14
1.3 Chaos in general relativity: a review and critique	15
1.3.1 Chaos detection	16
1.3.2 Time coordinate subtleties	16
1.3.3 Genericity and the choice of initial conditions	17
1.3.4 Papapetrou spin parameter confusion	17
1.3.4.1 Chaos in this thesis	18
2 Dynamics of spinning test particles in Kerr spacetime	21
Abstract	21
2.1 Introduction	21
2.2 Spinning test particles	24

2.2.1	Papapetrou equations	24
2.2.1.1	Spin supplementary conditions	25
2.2.1.2	A reformulation of the equations	25
2.2.1.3	Range of validity	27
2.2.2	Comments on the spin parameter	27
2.2.2.1	Bounds on S for stellar objects	27
2.2.2.2	Tidal disruption	29
2.2.2.3	The $S = 1$ limit	30
2.2.3	Symmetries and the parameterization of initial conditions	31
2.3	Lyapunov exponents	32
2.3.1	General discussion of Lyapunov exponents	32
2.3.2	Numerical calculation of Lyapunov exponents	35
2.3.2.1	The algorithm in detail	35
2.3.2.2	Gram-Schmidt and Lyapunov exponents	36
2.4	Relativity and Papapetrou subtleties	37
2.4.1	Phase space norm	37
2.4.2	Constraint complications	37
2.4.2.1	Constraint-satisfying deviations	38
2.4.2.2	A modified Gram-Schmidt algorithm	38
2.4.2.3	Two rigorous techniques	39
2.5	Implementation details	40
2.5.1	Some numerical comments	40
2.5.2	The Jacobian matrix	42
2.6	Integrability and chaos	44
2.6.1	Phase space and constants of the motion	44
2.6.2	Hamiltonian systems	45
2.6.2.1	Lyapunov exponents for Hamiltonian flows	45
2.6.2.2	Exponents for spinning test particles	45
2.7	Results	46
2.7.1	Chaos for $S = 1$	46
2.7.1.1	Maximally spinning Kerr spacetime	46
2.7.1.2	Schwarzschild spacetime revisited	51
2.7.1.3	Kerr and Schwarzschild compared	51
2.7.2	Dependence on S	54
2.7.3	Physically realistic spins	56
2.7.3.1	Vanishing Lyapunov exponents	56

2.7.3.2	Spin-induced phase differences	57
2.7.4	Comments on time, rescaling, and norms	59
2.8	Conclusions	60
	Acknowledgements	61
	Appendix: Full Jacobian	62
3	A survey of spinning test particle orbits in Kerr spacetime	66
	Abstract	66
3.1	Introduction	66
3.2	Spinning test particles	68
3.2.1	Equations of motion	68
3.2.2	General constraints	69
3.2.3	Kerr constraints	69
3.2.4	The spin parameter S	69
3.3	Parameterizing initial conditions	70
3.3.1	Energy and angular momentum parameterization	70
3.3.2	Orbital geometry parameterization	71
3.3.2.1	Kerr geodesics	72
3.3.2.2	From Kerr geodesics to Papapetrou initial conditions	74
3.3.2.3	Empirical orbital parameters	76
3.4	Lyapunov exponents	77
3.4.1	The principal exponent	77
3.4.2	The Papapetrou case	80
3.4.3	Chaos detector	82
3.4.4	Implementation notes	85
3.5	Results	85
3.5.1	Varying pericenter and orbital inclination	89
3.5.2	Varying eccentricity	92
3.5.3	Varying Kerr parameter a	93
3.5.4	Spin cutoffs for chaos	95
3.5.5	Retrograde orbits	95
3.5.6	Varying spin inclination	99
3.5.7	Deep integrations	101
3.5.8	$S = 0$ and chaos mimics	103
3.6	Conclusions	106
	Acknowledgements	107

Appendix: Spin vector formulation	107
4 Dynamics of binary black holes using the post-Newtonian approximation	129
Abstract	129
4.1 Introduction	129
4.2 The post-Newtonian equations of motion	131
4.2.1 The Hamiltonian formulation	131
4.2.2 Conserved quantities	135
4.2.2.1 Quantities conserved at all orders	135
4.2.2.2 Quantities conserved only through spin-orbit coupling	135
4.2.3 Numerical implementation	135
4.3 Parameterizing post-Newtonian orbits	136
4.3.1 Eccentric orbits	136
4.3.2 Quasi-circular orbits	138
4.3.3 Post-Newtonian orbits at various orders	139
4.4 The extreme mass-ratio case	139
4.4.1 Deviation from “geodesic” orbits	145
4.4.2 Comparing post-Newtonian and Papapetrou orbits	146
4.5 Investigating chaos in the post-Newtonian equations	149
4.5.1 A brief discussion of Lyapunov exponents	149
4.5.2 A survey of quasi-circular orbits	152
4.5.2.1 Varying spin directions	152
4.5.2.2 Varying initial frequencies	157
4.5.2.3 Varying spin magnitudes	158
4.5.2.4 Varying the PN terms	163
4.6 Conclusions and future work	170
Acknowledgements	171
5 Lyapunov exponents in constrained and unconstrained ODEs	174
Abstract	174
5.1 Introduction	174
5.2 Lyapunov exponents in unconstrained flows	176
5.2.1 The deviation vector method	177
5.2.2 The Jacobian method	179
5.2.2.1 Jacobian diagnostic	180
5.2.2.2 The principal exponent	180
5.2.2.3 Ellipsoids and multiple exponents	182

5.2.2.4	The algorithm in detail	185
5.2.3	The value of multiple exponents	187
5.2.4	Examples	188
5.2.4.1	The Lorenz system	188
5.2.4.2	The forced damped pendulum	190
5.3	Lyapunov exponents in constrained flows	192
5.3.1	Constraint complications	194
5.3.2	Constrained deviation vectors	194
5.3.2.1	Rescaling for constrained systems	195
5.3.2.2	A Jacobian method for the largest exponent	196
5.3.2.3	Ellipsoid constraint complications	196
5.3.3	Restricted Jacobian method	197
5.3.4	Constrained ellipsoid method	200
5.4	Comparing the methods	202
5.4.1	Speed	205
5.4.2	Robustness	206
5.5	Summary and conclusion	209
	Acknowledgements	209
	Appendix: Ellipsoid axes and the singular value decomposition	209
Afterword		214
	Lessons and advice	214
	A cautionary tale	216

List of Figures

2.1	The maximum spin angular momentum vs. mass for a rigidly rotating white dwarf . .	29
2.2	The maximum spin per unit mass s_{\max}/μ vs. mass for a white dwarf	30
2.3	White dwarf tidal disruption	31
2.4	The Jacobian matrix maps a tangent space ball to an ellipsoid	34
2.5	The orbit of a non-spinning test particle in maximal Kerr spacetime	47
2.6	The orbit of a maximally spinning ($S = 1$) particle in maximal Kerr spacetime	48
2.7	Lyapunov plot for an $S = 1$ particle in Kerr spacetime	49
2.8	Positive exponents and exponents indistinguishable from zero	50
2.9	The orbit of a maximally spinning ($S = 1$) test particle in Schwarzschild spacetime . .	52
2.10	Lyapunov plot in Schwarzschild spacetime with $S = 1$	53
2.11	Comparison of maximally spinning ($S = 1$) particle orbits in Kerr and Schwarzschild spacetime	54
2.12	Variation of largest Lyapunov exponent vs. S	55
2.13	Two orbits from the “bump” in Fig. 2.12(a)	56
2.14	The variation of $\lambda_{\max} M$ with final integration time	57
2.15	Lyapunov plots for proper time and coordinate time	59
2.16	Lyapunov plots for different rescaling times T	60
2.17	Lyapunov plots for two different norms	61
3.1	Rigorous Jacobian method compared to unrescaled deviation vector method	80
3.2	A comparison of chaotic and nonchaotic initial conditions	83
3.3	A chaos mimic	84
3.4	δr for a chaos mimic	84
3.5	δr for a chaotic orbit	85
3.6	The orbit of a maximally spinning ($S = 1$) test particle in maximal Kerr spacetime . .	86
3.7	Shading/coloring of parameter space for an $S = 1$ orbit	87
3.8	Shading/coloring of parameter space for an $S = 1$ orbit, alternate parameterization method	88
3.9	r_p - t map	90

3.10	r_p - t map, alternate parameterization method	91
3.11	Lyapunov plot for $e = 0.6$ chaos	92
3.12	Lyapunov plot for $e = 0.6$ chaos: verification by the Jacobian method	93
3.13	Spin cutoff map for $a = 1$ and $e = 0.5$	94
3.14	Spin cutoff map for $a = 1$ and $e = 0.01$	96
3.15	Spin cutoff map for $a = 0.5$ and $e = 0.5$	97
3.16	r_p - t map of retrograde orbits	98
3.17	Lyapunov exponents for varying values of initial spin components	100
3.18	Scatter plot of empirical pericenters $r_{p,P}$ vs. Lyapunov exponent for varying spin inclinations	100
3.19	Approximate Lyapunov exponent vs. time for a nonchaotic deep integration ($S = 0.1$)	102
3.20	Approximate Lyapunov exponent vs. time for a nonchaotic deep integration ($S = 10^{-4}$)	102
3.21	Approximate Lyapunov exponent vs. time for a chaotic deep integration	103
3.22	The orbit of a chaos mimic	104
3.23	Another chaos mimic	105
3.24	A chaos mimic Lyapunov plot	106
3.25	r_p - t map for $S = 0.9$, $a = 1$, and $e = 0.5$	109
3.26	r_p - t map for $S = 0.5$, $a = 1$, and $e = 0.5$	110
3.27	r_p - t map for $S = 0.2$, $a = 1$, and $e = 0.5$	111
3.28	r_p - t map for $S = 0.1$, $a = 1$, and $e = 0.5$	112
3.29	r_p - t map for $S = 10^{-4}$, $a = 1$, and $e = 0.5$	113
3.30	r_p - t map for $S = 1$, $a = 1$, and $e = 0.6$	114
3.31	r_p - t map for $S = 0.5$, $a = 1$, and $e = 0.6$	115
3.32	r_p - t map for $S = 0.1$, $a = 1$, and $e = 0.6$	116
3.33	r_p - t map for $S = 10^{-4}$, $a = 1$, and $e = 0.6$	117
3.34	r_p - t map for $S = 1$, $a = 1$, and $e = 0.01$	118
3.35	r_p - t map for $S = 0.5$, $a = 1$, and $e = 0.01$	119
3.36	r_p - t map for $S = 0.1$, $a = 1$, and $e = 0.01$	120
3.37	r_p - t map for $S = 10^{-4}$, $a = 1$, and $e = 0.01$	121
3.38	r_p - t map for $S = 1$, $a = 0.9$, and $e = 0.5$	122
3.39	r_p - t map for $S = 1$, $a = 0.7$, and $e = 0.5$	123
3.40	r_p - t map for $S = 1$, $a = 0.5$, and $e = 0.5$	124
3.41	r_p - t map for $S = 1$, $a = 0.4$, and $e = 0.5$	125
3.42	r_p - t map for $S = 1$, $a = 0.2$, and $e = 0.5$	126
3.43	r_p - t map for $S = 1$, $a = 0$, and $e = 0.5$	127

4.1	The orbit of two maximally spinning $10 M_{\odot}$ black holes	132
4.2	A highly eccentric orbit for two maximally spinning $10 M_{\odot}$ black holes	137
4.3	A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with Newtonian, 1PN, and 2PN terms turned on	140
4.4	A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with terms through spin-orbit coupling	141
4.5	A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with all PN terms present	142
4.6	A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with all PN terms present	143
4.7	A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with all PN terms present	144
4.8	PN and Papapetrou orbits compared	147
4.9	The error in the Papapetrou radial period $T_{r,P}$ vs. mass ratio for an eccentric orbit .	148
4.10	The error in the Papapetrou radial period $T_{r,P}$ vs. mass ratio for an near-circular orbit	149
4.11	Chaotic orbit of two maximally spinning $10 M_{\odot}$ black holes	150
4.12	Lyapunov plot for a chaotic PN orbit	150
4.13	Lyapunov exponents for 500 quasi-circular orbits as a function of total angular mo- mentum J for the $(10+10)M_{\odot}$ configuration at 240 Hz.	153
4.14	Lyapunov exponents for 500 quasi-circular orbits as a function of total angular mo- mentum J for the $(20+10)M_{\odot}$ configuration at 240 Hz	153
4.15	Lyapunov exponents for 500 quasi-circular orbits as a function of total angular mo- mentum J for the $(10+10)M_{\odot}$ configuration at 240 Hz	154
4.16	Lyapunov plot for the strongest $(10+10)M_{\odot}$ chaos	156
4.17	Angular distribution of chaos	156
4.18	Lyapunov exponents as a function of gravitational-wave frequency for quasi-circular orbits of two maximally spinning $10 M_{\odot}$ black holes	157
4.19	Lyapunov plots for frequencies on opposite sides of the transition to chaos, $(10+10)M_{\odot}$ configuration	158
4.20	Nonchaotic quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, $f_{\text{GW}} =$ 158 Hz	159
4.21	Chaotic quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, $f_{\text{GW}} = 160$ Hz	159
4.22	Lyapunov exponents as a function of gravitational-wave frequency for quasi-circular orbits of two maximally spinning black holes, $(20+10)M_{\odot}$ configuration	160
4.23	Lyapunov plots for frequencies on opposite sides of the transition to chaos, $(20+10)M_{\odot}$ configuration	160

4.24	The nonchaotic orbit of two maximally spinning black holes, $(20 + 10)M_{\odot}$ configuration	161
4.25	The chaotic orbit of two maximally spinning black holes, $(20 + 10)M_{\odot}$ configuration	162
4.26	Lyapunov exponents as a function of spin for quasi-circular orbits of two $10 M_{\odot}$ black holes	164
4.27	Lyapunov exponents as a function of spin for quasi-circular orbits of two black holes, $(20 + 10)M_{\odot}$ configuration	164
4.28	Lyapunov plots for spins on opposite sides of the transition to chaos, $(10 + 10)M_{\odot}$ configuration	165
4.29	Lyapunov plots for spins on opposite sides of the transition to chaos, $(20 + 10)M_{\odot}$ configuration	165
4.30	Nonchaotic quasi-circular orbit of two $10 M_{\odot}$ black holes, $S_2 = 0.83$	166
4.31	Chaotic quasi-circular orbit of two $10 M_{\odot}$ black holes, $S_2 = 0.86$	167
4.32	Nonchaotic quasi-circular orbit of two black holes, $(20 + 10)M_{\odot}$ configuration, $S_2 = 0.40$	168
4.33	Chaotic quasi-circular orbit of two black holes, $(20 + 10)M_{\odot}$ configuration, $S_2 = 0.58$	169
4.34	Lyapunov plot for the strongest chaos found in the $(20 + 10)M_{\odot}$ configuration	170
5.1	The Lorenz attractor	177
5.2	Comparison of the unrescaled and rescaled deviation vector methods for calculating the principal Lyapunov exponent of the Lorenz system	179
5.3	The natural logarithm of the tangent vector length vs. t for the Lorenz system	181
5.4	The Lorenz system with an evolving ellipsoid	182
5.5	Lyapunov plots for all ellipsoid axes in the Lorenz system	183
5.6	Lyapunov plots for two largest ellipsoid axes in the Lorenz system	186
5.7	θ vs. t for the forced damped pendulum	190
5.8	Time- 2π stroboscopic map for the forced damped pendulum	191
5.9	Lyapunov plots for all ellipsoid axes in the forced damped pendulum	192
5.10	Lyapunov plot for a simple pendulum	193
5.11	Comparison of the unrescaled and rescaled constrained deviation vector methods for calculating the principal Lyapunov exponent of the constrained forced damped pendulum	195
5.12	The natural logarithm of the tangent vector length vs. t for the constrained forced damped pendulum	196
5.13	Lyapunov plot for the constrained forced damped pendulum, restricted Jacobian method	198
5.14	The orbit of a spinning relativistic particle	199
5.15	Lyapunov plot for a spinning relativistic binary	200
5.16	Lyapunov plot for the constrained forced damped pendulum, constrained ellipsoid method	201

5.17	The orbit of a small spinning compact object in the spacetime of a rotating supermassive black hole	203
5.18	Lyapunov plot for a small spinning object	204
5.19	Lyapunov plots for all Lyapunov methods in the constrained forced damped pendulum	204
5.20	Lyapunov plot for the Lorenz system using the Gram-Schmidt algorithm, with the axes rescaled every $T = 10^{-3}$	207
5.21	Lyapunov plots in the Lorenz system for different rescaling times	208
5.22	Lyapunov plots for varying initial deviation vector size	208
5.23	Lyapunov plot for the Lorenz system using the singular value decomposition, frequent rescaling	211
5.24	Lyapunov plot for the Lorenz system using the singular value decomposition, no rescaling	211

List of Tables

2.1	Lyapunov exponents in Kerr spacetime	51
2.2	Lyapunov exponents in Schwarzschild spacetime	51
2.3	Phase shifts for $a = 0.5$ and $S = 10^{-5}$	58
2.4	Phase shifts for $a = 1$ and $S = 10^{-5}$	58
3.1	Local spin inclination angles θ_{local}	101
4.1	PN phase shifts for $e = 0.5$	145
4.2	PN phase shifts for $e = 0.05$	145
4.3	The prevalence of chaos in post-Newtonian quasi-circular orbits at 240 Hz	155
4.4	The prevalence of chaos in post-Newtonian quasi-circular orbits at 40 Hz	163
5.1	Comparison of different Lyapunov exponent methods applied to the forced damped pendulum	205
5.2	Timing comparison for different Lyapunov exponent methods applied to the forced damped pendulum	206

Chapter 1

Introduction

General relativity is nonlinear. As a result, in the solutions to general relativistic equations of motion lies the possibility of an extreme sensitivity on initial conditions—that is, general relativity is potentially chaotic. In classical physics, exactly solvable systems—such as the Kepler problem in Newtonian gravity—deservedly receive a great deal of attention, and so it is in relativity as well: the exactly solvable motion of a test particle orbiting a rotating black hole stands as a particularly remarkable example. But just as classical physics produces the chaotic orbit of Pluto [1] and the chaotic obliquity of Mars [2], so must general relativity ultimately produce chaos as well. It is the aim of this thesis to explore the dynamics of a particular family of potentially chaotic relativistic systems: binary systems composed of spinning compact objects, such as white dwarfs, neutron stars, and black holes. Compact binaries are promising sources of gravitational waves, both for ground-based detectors (such as LIGO, the Laser Interferometer Gravitational-wave Observatory [3]) and space-based detectors (such as the proposed LISA project, the Laser Interferometer Space Antenna [4]). Chaos in these systems would have significant observational consequences; in particular, chaos could complicate attempts to detect gravitational waves from these sources. This is because current *matched filter* detection methods rely on covering the parameter space with a discrete mesh, with each choice of mesh size achieving a particular signal-to-noise ratio [5–7]. Chaos in the gravitational wave sources would cause the number of such parameter-space templates needed to fit a given signal to increase exponentially with decreasing mesh size [8], potentially leading to a computationally intractable number of templates. It is this concern, and the intrinsic interest of understanding the nonlinear dynamics of general relativity, that motivate the present work.

Chapters 2 and 3 of this thesis are concerned with the dynamics of *extreme mass-ratio binaries*, where the larger object is a rotating (Kerr) black hole. If the smaller body has mass μ , and the central body has mass M , then we require that $\mu \ll M$, so that the mass of the small body does not affect the background spacetime.¹ In the extreme mass-ratio limit, which is well-satisfied by, e.g., compact objects orbiting supermassive black holes in galactic nuclei, the general relativistic

¹The gravitational radiation from these systems should be detectable by LISA.

equations of motion simplify considerably, becoming a set of ordinary differential equations—much simpler than their cousins, the partial differential equations of full general relativity.² In the case of a *nonspinning* small particle, the equations simplify *too* much from our perspective: a nonspinning test particle follows Kerr geodesics, which are *fully integrable*, i.e., there are as many constants of the motion as there are coordinates. As a consequence, there is no possibility of chaos. To restore this possibility, we add spin to the smaller body, which introduces new (spin) degrees of freedom while simultaneously destroying one of the constants. We discuss this system, described by the *Papapetrou equations*, in Sec. 1.1.1.

In Chapter 4, we consider the dynamics of spinning compact binaries using the *post-Newtonian (PN) equations* with spin [9–11], which are essentially a perturbative expansion of full general relativity in v/c ; these equations allow us to consider the case of *comparable masses*. We use a Hamiltonian formulation that is particularly well-suited to the study of the PN system as a nonlinear dynamical system. The post-Newtonian Hamiltonian we use contains several spin interaction terms, including the traditional spin-orbit coupling and a spin-spin coupling term, as well as mass monopole/spin-induced quadrupole interaction terms that are of the same order as the spin-spin coupling [9]. Since these quadrupole terms are too large to neglect, we specialize in the present case to binary black holes, for which the quadrupole terms are known exactly. (In neutron stars these terms depend on the poorly known equation of state.³) We pay particularly close attention to *quasi-circular* orbits (“as circular as possible,” given nonvanishing spins). These are particularly relevant for gravitational wave generation and detection, since many binary systems are expected to circularize due to gravitational radiation reaction [13] (and, moreover, stay circular [14]). In Chapter 4, we also consider the extreme mass-ratio limit of the PN equations, in order to make contact with the Papapetrou approximation considered in Chapters 2 and 3.

For both the Papapetrou and the post-Newtonian systems, our primary method for quantifying chaos is the calculation of *Lyapunov exponents*, which measure the average divergence rate of nearby phase-space trajectories. A Lyapunov exponent greater than zero indicates that nearby trajectories separate exponentially fast, which is a characteristic of chaotic systems. (See Chapter 2 for further discussion of the technical requirements for chaos.) In the process of calculating these exponents in the Papapetrou equations, we encountered several difficulties. The Lyapunov exponents can be calculated in several different ways, but for the Papapetrou system we could not initially get these methods to agree. Furthermore, even after this discrepancy was resolved,⁴ it became clear that the constrained nature of the Papapetrou equations (which enforce, for example, normalization conditions on the spin and 4-momentum) leads to significant complications in calculating the characteristic exponents. Addressing these difficulties led to the development of general techniques for dealing with

²Hats off to the brave souls of numerical relativity who tackle the partial differential equations directly.

³An approximate treatment of neutron star binaries is still possible [12], and is left for future work.

⁴See the Afterword for an account of the discrepancy and its resolution.

chaos in constrained dynamical systems, and eventually resulted in a suite of computational tools for solving the problem in a general setting. These techniques, and their software implementation, are discussed in Chapter 5; they allowed most of the results in Chapter 4 to be obtained rather quickly and painlessly.⁵ Chapter 5 is logically independent of Chapters 2–4; readers of this thesis⁶ less familiar with dynamical systems theory than general relativity may wish to refer forward to Chapter 5 while reading Chapters 2–4.

It is important to note (as mentioned briefly above) that we ignore gravitational radiation reaction throughout this thesis. The two systems we consider—the Papapetrou equations and the post-Newtonian equations—are conservative, Hamiltonian dynamical systems. The addition of dissipative radiation, and the resulting inevitable plunge, would destroy the formal possibility of chaos, since the mathematical definition of chaos relies on an infinite-time limit (Sec. 1.2). For the purposes of gravitational-wave templates, we may ignore radiation reaction, since any chaos present in the conservative limit, if it occurs on a short-enough timescale, will manifest itself in the dissipative dynamics as well. As discussed in Sec. 1.2 below, the Lyapunov exponent (through its inverse) provides exactly the timescale needed to make this judgment.

In this thesis, we work almost exclusively in geometric units, which set $G = c = 1$, so that length, time, and momentum have units of mass, while energy and angular momentum have units of mass squared. We denote relativistic 4-vectors either with a Greek superscript (e.g., v^μ , $\mu \in \{0, 1, 2, 3\}$) or with a vector arrow (e.g., \vec{v}). We set Euclidean vectors in boldface (e.g., $\boldsymbol{\xi}$). The symbol \log denotes in all cases the natural logarithm; on the rare occasion that base-10 logarithms are needed, we denote them by \log_{10} .⁷

Finally, we note that the chapters in this thesis are reasonably self-contained. This is convenient for the author, since the chapters represent lightly edited versions of papers prepared for separate publication. It is also convenient for the reader, since those interested primarily in, e.g., a survey of chaos in Kerr spacetime can skip directly to Chapter 3. The price we pay for this modularity is a minor amount of redundancy, but even this is a feature, not a bug: the topics receiving the most coverage constitute precisely the foundational material essential for understanding the primary results of the thesis. This recalls the so-called “spiral approach” to pedagogy, wherein the same topic is covered several times from different points of view, often with increasing levels of sophistication. The treatment of Lyapunov exponents is an example, with a brief introduction in Sec. 1.2, a longer discussion in Chapters 2 and 3, and then a thorough examination in Chapter 5.

⁵Obtaining the results in Chapters 2 and 3 was painful.

⁶Assuming there are any.

⁷We use “log” for base- e logarithms (in place of “ln,” which is pronounced “log” half the time anyway) because Nature does not particularly care that we have ten fingers. (If only we had e fingers—we could then avoid all confusion on this point.) This usage is also an important reminder for the computationally inclined: in virtually all computer languages, the user will be unpleasantly surprised if his code expects, e.g., `log(10)` to return 1.

1.1 The Papapetrou and post-Newtonian equations

We give here a brief introduction to the principal systems studied in this thesis, both of which model, in their respective ranges of validity, relativistic binary systems consisting of spinning compact objects.

1.1.1 Papapetrou (or Dixon)

The Papapetrou equations [15] describe the motion of a spinning test particle in an arbitrary background spacetime. More specifically, the equations model a *pole-dipole* particle, which includes the spin dipole moment but ignores higher order effects (such as the mass quadrupole). They are also conservative, and ignore the dissipative effects of gravitational radiation. Because of its conceptual clarity, the formulation due to Dixon [16] is the starting point for most investigations, including ours. Dixon writes the equations of motion in terms of the 4-momentum p^α and spin tensor $S^{\alpha\beta}$, which are defined by integrals of the particle's stress-energy tensor $T^{\alpha\beta}$ over an arbitrary spacelike hypersurface Σ :

$$p^\alpha(\Sigma) = \int_{\Sigma} T^{\alpha\beta} d\Sigma_\beta \quad (1.1)$$

$$S^{\alpha\beta}(\vec{z}, \Sigma) = 2 \int_{\Sigma} (x^{[\alpha} - z^{[\alpha} T^{\beta]\gamma}) d\Sigma_\gamma, \quad (1.2)$$

where \vec{z} is the coordinate of the center of mass. (Since the stress-energy tensor has vanishing divergence, these definitions are independent of Σ [16].) The equations of motion for a spinning test particle are then

$$\begin{aligned} \frac{dx^\mu}{d\tau} &= v^\mu \\ \nabla_{\vec{v}} p^\mu &= -\frac{1}{2} R^\mu_{\nu\alpha\beta} v^\nu S^{\alpha\beta} \\ \nabla_{\vec{v}} S^{\mu\nu} &= 2p^{[\mu} v^{\nu]}, \end{aligned} \quad (1.3)$$

where v^μ is the 4-velocity, i.e., the tangent to the particle's worldline, $R^\mu_{\nu\alpha\beta}$ is the Riemann curvature tensor, and brackets denote antisymmetrization [so that $p^{[\mu} v^{\nu]} = \frac{1}{2}(p^\mu v^\nu - p^\nu v^\mu)$]. The 4-momentum equation in Eq. (1.3) includes a term involving a coupling of the spin to the Riemann curvature tensor of the spacetime. In the absence of spin (all components of $S^{\alpha\beta}$ zero), or in flat spacetime (all components of $R^\mu_{\nu\alpha\beta}$ zero), we recover the geodesic equation,

$$\nabla_{\vec{v}} \vec{p} = 0. \quad (1.4)$$

For nonzero spin or curvature, the 4-momentum deviates from geodesics due to the spin-curvature coupling.

For the rest of this thesis, whenever we refer to “the Papapetrou equations,” we almost always mean the system described by Eq. (1.3) with the background spacetime given by a Kerr black hole. An astrophysical Kerr black hole is fully specified by the mass M and the spin angular momentum per unit mass a .⁸ Geodesic motion of a particle around a Kerr black hole is constrained by three integrals of the motion: the energy E , the z angular momentum L_z , and the Carter constant Q . As noted above, these are sufficient to guarantee the absence of chaotic solutions for Kerr geodesics.⁹ For the non-geodesic motion represented by the Papapetrou equations, we use the metric for Kerr spacetime to determine the Riemann tensor appearing in Eq. (1.3). In this case, the Papapetrou equations model a spinning particle orbiting a (much larger) spinning black hole. (A “particle” in this context is simply a second compact object—a white dwarf, neutron star, or black hole—and should not be confused with an elementary particle.)

Two of the constants in Kerr geometry, the energy and axial angular momentum, are manifestations of a general phenomenon: to each symmetry of the metric (typically represented by a Killing vector) there corresponds a conserved quantity (which is essentially Nöther’s theorem). This result holds even in the presence of spin: for an arbitrary spacetime, each Killing vector $\vec{\xi}$ leads to a constant

$$C_{\xi} = \xi^{\mu} p_{\mu} - \frac{1}{2} \xi_{\mu;\nu} S^{\mu\nu}, \quad (1.5)$$

where $\xi_{\mu;\nu} = -\xi_{\nu;\mu}$ is the covariant derivative of the Killing vector. The time-translation symmetry of the Kerr metric then gives energy conservation,

$$-E = p_t - \frac{1}{2} g_{t\mu,\nu} S^{\mu\nu}, \quad (1.6)$$

while rotational symmetry about the black hole’s spin axis gives z angular momentum conservation,

$$J_z = p_{\phi} - \frac{1}{2} g_{\phi\mu,\nu} S^{\mu\nu}. \quad (1.7)$$

The Carter constant does not arise from an explicit symmetry of the metric, but rather from a Killing tensor $K_{\mu\nu}$ [17] (satisfying $K_{\mu(\nu;\alpha)} = 0$); the corresponding quantity is not conserved when the particle has nonvanishing spin, although it deviates from a constant only at linear order in the spin [18].

As discussed in Chapter 2, we rewrite the Papapetrou equations in terms of the more intuitive

⁸This is a special case of the celebrated “no-hair theorem” for black holes, which says that a black hole is specified completely by its mass, spin, and charge. For astrophysical black holes, charge is unimportant, since any charge would be shorted out at values too small to be gravitationally significant.

⁹Here we consider there to be three coordinates, say the Boyer-Lindquist coordinates (r, θ, ϕ) . If we include the time t on our list of coordinates, we add the conserved particle rest mass μ to our list of constants, and the conclusion is the same. This leaves mass and spin as the only relevant parameters.

spin vector,

$$S^\mu = -\frac{1}{2}\epsilon^\mu_{\nu\alpha\beta} u^\nu S^{\alpha\beta}, \quad (1.8)$$

where $u^\nu = p^\nu/\mu$, so that Eq. (1.3) can be reformulated as a system of 12 coupled nonlinear ordinary differential equations. These equations constitute the dynamical system that is the focus of our study in Chapters 2 and 3.

1.1.2 Post-Newtonian (with spin)

Whereas the Papapetrou equations represent a system in full general relativity, but are valid only in the test-particle limit, the post-Newtonian equations are valid only for $v \ll c$, but their range of validity includes even equal-mass binaries. In schematic form, the PN equations of motion read

$$\mu\ddot{\mathbf{r}} = \mathbf{F}_0 + \epsilon \mathbf{F}_1 + \epsilon^{3/2} \mathbf{F}_{3/2} + \epsilon^2 \mathbf{F}_2 + \epsilon^{5/2} \mathbf{F}_{5/2} + \dots, \quad (1.9)$$

where \mathbf{r} is the relative separation and $\epsilon = (v/c)^2 \sim M_{\text{total}}/r$. The zeroth term is classical Newtonian gravity, while the higher-order terms represent relativistic corrections. It is traditional to refer to the various terms by their order in $(v/c)^2$, so that the first term is 1PN, the second is 2PN, and so on. The $\frac{5}{2}$ PN term is the radiation reaction due to gravitational radiation, which we ignore throughout this thesis, since we are interested primarily in the long-time evolution of the system.¹⁰ The claim in [19] (and subsequent papers) that the PN equations have chaotic solutions was an early motivation for the work in this thesis. While we cannot (yet) comment definitively on these results, we do offer a broad survey of quasi-circular orbits, an important special case.

The Hamiltonian formulation we use for the post-Newtonian equations, as developed in [9], includes the following terms:

$$H = H_N + H_{\text{PN}} + H_{\text{SO}} + H_{\text{SS}}, \quad (1.10)$$

where we use subscripts to label the different contributions: Newtonian (N); traditional post-Newtonian through 2PN, without spin (PN); spin-orbit coupling (SO); and terms quadratic in the spins (SS). The full form of the terms appears in Chapter 4; for now, we simply note that the spin-orbit (SO) coupling is effectively a $\frac{3}{2}$ PN effect, and the terms quadratic in the spins (SS) are effectively 2PN effects [20]. The Hamiltonian describes an effective one-body system, and so is a function of the relative position \mathbf{X} and the relative momentum \mathbf{P} ; it also depends on the spins \mathbf{S}_1 and \mathbf{S}_2 of the two objects. The equations of motion can be derived using Poisson brackets:

$$\frac{d\mathbf{X}}{dt} = \{\mathbf{X}, H\} = +\frac{\partial H}{\partial \mathbf{P}}, \quad \frac{d\mathbf{P}}{dt} = \{\mathbf{P}, H\} = -\frac{\partial H}{\partial \mathbf{X}} \quad (1.11)$$

¹⁰We plan to examine the effects of radiation reaction, as well as the recently calculated 3PN corrections, in future work.

for the position and conjugate momentum, and

$$\frac{d\mathbf{S}_1}{dt} = \{\mathbf{S}_1, H\} = \frac{\partial H}{\partial \mathbf{S}_1} \times \mathbf{S}_1 \equiv \boldsymbol{\Omega}_1 \times \mathbf{S}_1 \quad (1.12)$$

and

$$\frac{d\mathbf{S}_2}{dt} = \{\mathbf{S}_2, H\} = \frac{\partial H}{\partial \mathbf{S}_2} \times \mathbf{S}_2 \equiv \boldsymbol{\Omega}_2 \times \mathbf{S}_2 \quad (1.13)$$

for the spins.¹¹ Since each vector has three components, these equations represent 12 coupled non-linear ordinary differential equations, which constitute the dynamical system studied in Chapter 4.

1.1.3 Normalized units and μM

We work in *normalized units* throughout much of this thesis. This involves measuring distances and times in terms of M and momenta in terms of μ . The meaning of M and μ depend on context; for the Kerr/Papapetrou case, M is the mass of the central black hole, while μ is the mass of the test particle. In the post-Newtonian case, M is the total mass $m_1 + m_2$, and μ is the reduced mass $m_1 m_2 / M$. When working in normalized units, we might refer to, for example, a particle at $r = 10$. This means that the particle is a distance $r = 10 M = 10 GM/c^2$ from the origin of the coordinate system.

The use of normalized units explains what seems to be a strange choice, employed in Chapters 2 and 3 and in [21]: the spin parameter S , which is the magnitude of the spin of the Papapetrou test particle, is measured in terms of the product μM , not (as one might expect) in terms of μ^2 . This choice of units is forced on us if (1) we measure lengths and times in terms of the mass M of the central body; (2) we measure momenta in terms of the mass μ of the small body; and (3) we wish to avoid introducing seemingly arbitrary factors of μ/M into the equations.

The reason for the strange choice of units can be motivated heuristically by writing the total angular momentum as the sum of the orbital angular momentum and the spin: $\mathbf{J} = \mathbf{L} + \mathbf{S}$. For simplicity, suppose that only the small body is spinning [as is the case when the central body is a nonrotating (Schwarzschild) black hole], and that the small body is a maximally spinning black hole, so that $\mathbf{S} = \mu^2 \hat{\mathbf{S}}$, where we use a hat to indicate a unit vector. Now, the orbital angular momentum is $\mathbf{L} = \mathbf{X} \times \mathbf{P}$, i.e, essentially a length times a momentum. To convert to normalized units, we need to factor out an M from \mathbf{X} and a μ from \mathbf{P} :

$$\mathbf{X} \times \mathbf{P} = \mu M (\tilde{\mathbf{X}} \times \tilde{\mathbf{P}}), \quad (1.14)$$

where we indicate vectors written in normalized units with tildes. In order to convert the total

¹¹This uses the canonical angular momentum Poisson bracket $\{S^i, S^j\} = \epsilon^{ijk} S^k$.

angular momentum to these units, we should divide by μM , so that

$$\tilde{\mathbf{J}} = \frac{\mathbf{J}}{\mu M} = \frac{\mathbf{L}}{\mu M} + \frac{\mathbf{S}}{\mu M} \equiv \tilde{\mathbf{L}} + \tilde{\mathbf{S}}. \quad (1.15)$$

where $\tilde{\mathbf{S}} = \mathbf{S}/\mu M$. If instead we insist on measuring the spin in terms of μ^2 , we have

$$\tilde{\mathbf{J}} = \frac{\mathbf{J}}{\mu M} = \frac{\mathbf{L}}{\mu M} + \frac{\mathbf{S}}{\mu M} = \tilde{\mathbf{L}} + \frac{\mu^2}{\mu M} \tilde{\mathbf{S}} = \tilde{\mathbf{L}} + \frac{\mu}{M} \tilde{\mathbf{S}}. \quad (1.16)$$

The explicit factor of μ/M is the price we pay for measuring S in units different from those used for \mathbf{J} and \mathbf{L} .

For the Papapetrou equations, the choice to measure S in terms of μM places sharp restrictions on the magnitude of \tilde{S} , as discussed at length in Chapter 2. Briefly, for a maximally spinning black hole, the maximum spin is μ^2 , so the maximum spin parameter is

$$\tilde{S}_{\max} = \frac{S}{\mu M} = \frac{\mu^2}{\mu M} = \frac{\mu}{M} \ll 1, \quad (1.17)$$

where in the final step we have imposed the test-particle condition $\mu \ll M$. Thus, the “spin parameter” appearing in the Papapetrou equations *must be small compared to unity*, at least for black holes. (We show in Chapter 2 that this result is true in general.) It is only in the case $\mu = M$ that $\tilde{S} = 1$; *mathematically*, the Papapetrou equations are perfectly valid in this limit, but they no longer realistically model any *physical* system. This point was not emphasized strongly in Suzuki and Maeda [21], leading some (and perhaps even the authors) to overstate the prevalence of chaos in Schwarzschild spacetime (see Sec. 1.3).

We note that the units discussed above pose no real restrictions on the magnitude of \tilde{S} for the post-Newtonian case, since the PN equations are valid for the case $\mu = M$, i.e., an equal-mass binary. The only difference is that (as in Chapter 4) we would write instead $m_1 = m_2$ for an equal-mass binary, reserving μ and M for the reduced and total mass, respectively.

For most of this thesis, we will work in normalized units, so for the most part we will omit tildes or any other special notation. The variable S , for example, means S measured in units of μM , so that $S = 1$ means that $S = 1 \mu M$. The only major exception is Chapter 4, where the spins we report are typically measured in more conventional units; e.g., $S_1 = 1$ means $S_1 = m_1^2$.

1.1.4 Normalization constraints

Both the Papapetrou equations and the PN equations possess some simple normalization constraints that are preserved by the equations of motion. The Papapetrou spin satisfies

$$S^\mu S_\mu = \frac{1}{2} S^{\mu\nu} S_{\mu\nu} = S^2, \quad (1.18)$$

while the PN spins satisfy $S_1^2 = \mathbf{S}_1 \cdot \mathbf{S}_1$ and $S_2^2 = \mathbf{S}_2 \cdot \mathbf{S}_2$, so that S , S_1 , and S_2 are constants of the motion. In the Papapetrou case, we have the additional requirement that the squared 4-momentum yield the (negative of) the particle's rest mass:

$$p^\mu p_\mu = -\mu^2, \quad (1.19)$$

which we typically write as

$$p^\mu p_\mu = -1 \quad (1.20)$$

in normalized units.

1.1.5 Spin supplementary conditions and the center of mass

The Papapetrou equations are under-determined as written in Eq. (1.3); they require a *spin supplementary condition* (SSC) to fix the center of mass of the spinning body [16]. The condition advocated and used in Chapter 2 (following [16]) is

$$p_\mu S^{\mu\nu} = 0. \quad (1.21)$$

This condition picks out a unique worldline, which we identify as the center of mass of the body. Indeed, if we define the “rest frame” of the body as the frame where the spatial part of the 4-momentum (that is, the conventional 3-momentum) vanishes, then Eq. (1.21), the definition of $S^{\alpha\beta}$ [Eq. (1.2)], and the condition $p^i = 0$ give the relativistic generalization for the Newtonian center of mass:

$$z^i = \frac{\int_{t=\text{const.}} x^i T^{00} d^3x}{\int_{t=\text{const.}} T^{00} d^3x}. \quad (1.22)$$

The choice of Eq. (1.21) for the supplementary condition is not the only one in the literature; the condition

$$v_\mu S^{\mu\nu} = 0 \quad (1.23)$$

is advocated by many authors (see [22] for a review). Because this condition is satisfied by a family of helical worldlines filling a cylinder with frame-dependent radius [16, 23]—and, not least, because condition (1.21) allows for an explicit relationship between v^μ and p^μ to be derived, which is especially convenient for numerical calculation—we follow Dixon in enforcing $p_\mu S^{\mu\nu} = 0$.

It is sometimes objected that it is pedantic to argue over which of Eq. (1.21) or Eq. (1.23) is “correct”: to the extent that the approximations represented by the Papapetrou equations are valid, the two conditions are identical. This objection is essentially correct; the two conditions differ at third order in the spin (Chapter 2), so the difference between the two conditions is negligible for physically realistic spins. Perhaps surprisingly, though, the general subject of spin supplementary conditions

is not pedantic—when considering the post-Newtonian limit, it becomes clear that conditions (1.21) and (1.23), the subject of so much argument, are not only identical, but also *disagree* with the traditional post-Newtonian center of mass.

To see the problem with the standard covariant SSCs [Eqs. (1.21) and (1.23)], consider the PN Hamiltonian for the spin-orbit coupling (discussed in Chapter 4):

$$H_{\text{SO}} = \frac{\mathbf{L} \cdot \mathbf{S}_{\text{eff}}}{r^3}, \quad (1.24)$$

where \mathbf{L} is the angular momentum, r is the relative separation, and

$$\mathbf{S}_{\text{eff}} = \left(2 + \frac{3}{2} \frac{m_2}{m_1}\right) \mathbf{S}_1 + \left(2 + \frac{3}{2} \frac{m_1}{m_2}\right) \mathbf{S}_2. \quad (1.25)$$

For the test particle limit to apply, one body must be much heavier than the other, say $m_1 \gg m_2$. In this limit, to lowest order the contribution of the lighter body to the spin-orbit coupling is

$$H_{\text{SO}} = \frac{3}{2} \frac{\mathbf{L} \cdot \mathbf{S}_2}{r^3} \frac{m_1}{m_2}. \quad (1.26)$$

In the limit $m_1 \gg m_2$, we can write $M \approx m_1$ for the total mass; the force derived from Eq. (1.26) is then

$$-\frac{3M}{r^3} \left(\mathbf{v} \times \mathbf{S} - \frac{3}{2} (\hat{\mathbf{n}} \cdot \mathbf{v}) (\hat{\mathbf{n}} \times \mathbf{S}) - \frac{3}{2} \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot [\mathbf{v} \times \mathbf{S}]) \right), \quad (1.27)$$

where $\hat{\mathbf{n}} = \mathbf{r}/r$ and we write \mathbf{S} in place of \mathbf{S}_2 for the spin of the small body. On the other hand, if we expand the spin-curvature coupling in the Papapetrou equations [Eq. (1.3)] for a Kerr background in v/c , then to the same order (1PN) as the spin-orbit coupling we have

$$-\frac{1}{2} R^\mu_{\nu\alpha\beta} v^\nu S^{\alpha\beta} = -\frac{3M}{r^3} (\mathbf{v} \times \mathbf{S} - (\hat{\mathbf{n}} \cdot \mathbf{v}) (\hat{\mathbf{n}} \times \mathbf{S}) - 2 \hat{\mathbf{n}} (\hat{\mathbf{n}} \cdot [\mathbf{v} \times \mathbf{S}])), \quad (1.28)$$

in disagreement with Eq. (1.27).

The resolution to this discrepancy is given explicitly in [24], and it was implicit long before in [25]: the forces in Eq. (1.27) and Eq. (1.28) are *both* correct, but they refer to *different locations of the center of mass*. The “natural” PN center of mass is defined by an integral involving the total mass, kinetic energy, internal energy, and gravitational potential energy (as given explicitly in [20]), which differs from the Papapetrou center of mass at 1PN order:

$$\mathbf{z}_{\text{PN}} = \mathbf{z}_{\text{Papa}} + \frac{1}{2\mu} \mathbf{v} \times \mathbf{S}, \quad (1.29)$$

so that the relative separation vectors satisfy

$$\mathbf{r}_{\text{PN}} = \mathbf{r}_{\text{Papa}} + \frac{1}{2\mu} \mathbf{v} \times \mathbf{S}, \quad (1.30)$$

where we write μ for the mass of the smaller body. Plugging \mathbf{r}_{Papa} from Eq. (1.30) into Eq. (1.28) yields Eq. (1.27).

As a final comment, we note that the difference between the PN and Papapetrou CM coordinates is small in all cases of physical interest, so that the orbits always virtually coincide—in particular, conclusions regarding chaos are not sensitive to the choice of CM coordinate. If we use normalized units (distances in terms of M , momenta in terms of μ , S in terms of μM), then Eq. (1.30) gives

$$|\mathbf{r}_{\text{PN}} - \mathbf{r}_{\text{Papa}}| = \left| \frac{1}{2} \mathbf{v} \times \mathbf{S} \right| = \left| \frac{1}{2} \mathbf{v} \times \mathbf{S} \right| \leq \frac{vS}{2}, \quad (1.31)$$

which is very small, since both v and S (when measured in terms of μM) are much smaller than unity. Of course, Eq. (1.30) requires the velocity to satisfy $v \ll 1$, but the result holds even when $v \sim 1$. In this case, the appropriate generalization of Eq. (1.30), in terms of the Lorentz factor $\gamma = (1 - v^2)^{-1/2}$, is [26]

$$\mathbf{r}_{\text{PN}} = \mathbf{r}_{\text{Papa}} + \frac{\gamma}{\mu(\gamma + 1)} \mathbf{v} \times \mathbf{S}, \quad (1.32)$$

so that (again in normalized units)

$$|\mathbf{r}_{\text{PN}} - \mathbf{r}_{\text{Papa}}| = \left| \frac{\gamma}{\gamma + 1} \mathbf{v} \times \mathbf{S} \right| \leq vS \leq S \ll 1. \quad (1.33)$$

1.1.6 Degrees of freedom

The constraints in our two model systems—both of which consist of 12 coupled differential equations—reduce the true number of degrees of freedom, and hence (Sec. 1.2 below) the number of Lyapunov exponents. The Papapetrou system has two normalization constraints [Eqs. (1.18) and (1.20)] and an orthogonality constraint [the spin supplementary condition, Eq. (1.21)], each of which eliminates one degree of freedom. Since time is not a dynamical variable in this context, we eliminate one more degree of freedom by projecting onto a spacelike hypersurface, leaving 8 degrees of freedom. For the post-Newtonian case, the two spin normalization constraints reduce the 12 apparent degrees of freedom to 10. In the extreme mass-ratio limit of the post-Newtonian system, one of the spins is fixed in magnitude and direction, eliminating two more degrees of freedom, leaving 8 total—and thus corresponding exactly to the Papapetrou case.

1.2 Lyapunov exponents and chaos

The Lyapunov exponents of a dynamical system provide a quantitative measure of the separation rate of nearby trajectories in phase space. A particularly intuitive method for visualizing these exponents, developed in detail throughout this thesis, is to imagine the time-evolution of an infinitesimal ball centered on a particular initial condition. Under the action of the flow, this ball is deformed into an ellipsoid—stretched in some directions, squeezed in others. In general, the (semi)axes r_i of the ellipsoid grow geometrically (that is, exponentially), so that on average $r_i(t) = L_i^t$ for some constant L_i . These L_i are the *Lyapunov numbers*, and the Lyapunov exponents are then

$$\lambda_i = \log L_i = \lim_{t \rightarrow \infty} \frac{\log [r_i(t)]}{t}. \quad (1.34)$$

(The infinite-time limit is part of the formal definition, but of course in practice we must use a finite cutoff; see Sec. 1.2.5.) For a system with system with n degrees of freedom, the initial n -dimensional ball deforms into an n -dimensional ellipsoid, yielding n Lyapunov exponents.

1.2.1 Deviation vector method

A less formal method, which makes contact with the intuitive notion of the “exponential separation” of nearby trajectories, can be used to extract the *largest* Lyapunov exponent. Heuristically, if we consider two initial conditions separated by a small deviation vector of length ϵ_0 , then in chaotic systems the separation (by definition) grows exponentially with time:

$$\epsilon(t) = \epsilon_0 e^{\lambda t}. \quad (1.35)$$

Since a generic deviation vector must have a nonzero component in the direction of greatest stretching, eventually it points essentially parallel to the longest ellipsoid axis, so the separation rate λ in Eq. (1.35) is the largest Lyapunov exponent. Its inverse is a time: $t_\lambda = 1/\lambda$, which we refer to as the *Lyapunov timescale* corresponding to λ ; nearby initial conditions diverge by a factor of e in a time t_λ . If the largest Lyapunov exponent vanishes, then the system is not chaotic.

Translating the informal definition of exponential divergence in Eq. (1.35) into a concrete algorithm, though not as straightforward as it seems, is not difficult: by choosing a deviation vector, and then monitoring its length as a function of time, we can estimate the largest Lyapunov exponent using

$$\lambda = \frac{\log [r(t)]}{t}, \quad (1.36)$$

where $r(t) = \epsilon/\epsilon_0$ is the length of the deviation after time t . We mention some subtleties and caveats that apply to this method throughout this thesis, especially in Chapter 5. We refer to this

(relatively) straightforward method as the *deviation vector method*.

1.2.2 Jacobian method

We can sharpen our understanding of Lyapunov exponents by making more precise the meaning of a “deviation vector.” Consider an autonomous system of first-order ordinary differential equations:

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}), \quad (1.37)$$

where $\mathbf{y} = \mathbf{y}(x_1, \dots, x_n)$. The requirement of *first-order* equations forces us to consider the “coordinates” $\{x_1, \dots, x_n\}$ to be the full phase space, so that in a typical mechanical system a “point” is a list of positions and velocities. Writing $\delta\mathbf{y}$ for a small deviation vector in this phase space, we can express the distance between the two trajectories by using the best linear approximation of \mathbf{f} ,

$$\frac{d(\delta\mathbf{y})}{dt} = \mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) = \mathbf{Df} \cdot \delta\mathbf{y}, \quad (1.38)$$

to lowest order in $\delta\mathbf{y}$, where \mathbf{Df} is the Jacobian matrix of the system $[(\mathbf{Df})_{ij} = \partial f_i / \partial x^j]$. By considering a complete set of deviation vectors spanning an infinitesimal ball, and following their evolution into an ellipsoid, we can extract all n exponents of the system, using an algorithm discussed in Chapters 2 and 5. Since the numerical implementation of this technique requires the Jacobian matrix, we refer to it as the *Jacobian method*.

1.2.3 The origin of exponential separation

We take an opportunity here to examine a fundamental question: Why do so many dynamical systems exhibit exponential separation of nearby initial conditions? As a first step, consider the one-dimensional version of Eq. (1.38):

$$\frac{d(\delta y)}{dt} = f'(x) \delta y, \quad (1.39)$$

which is an equation of motion for the deviation. At any particular point, $f'(x)$ simply evaluates to a number, so Eq. (1.39) says that at each point $d(\delta y)/dt \propto \delta y$. This is precisely the equation for exponential growth: *locally*, at least, the deviation grows exponentially. Of course, if $f'(x) < 0$, then this “growth” is really a decay, and so on average the “growth” might actually be zero. [In fact, for one- or two-dimensional dynamical systems (described by autonomous differential equations), this must happen: the growth, and hence the Lyapunov exponent(s), must in general average out to zero.¹²] If, on the other hand, the average growth is positive, it will be exponential, by virtue of the

¹²This follows from the Poincaré-Bendixon theorem [27].

equation of motion for the deviation [the multidimensional generalization of Eq. (1.39)],

$$\frac{d(\delta \mathbf{y})}{dt} = \mathbf{Df} \cdot \delta \mathbf{y}. \quad (1.40)$$

Thus, we see that the origin of exponential growth is the *linear* nature of the local approximation to the system—that is, its differentiability.

1.2.4 Constrained systems

The standard methods for the numerical calculation of Lyapunov exponents apply to *unconstrained* systems of differential equations (as discussed in textbooks such as [27] and [28]), but the addition of constraints to the phase-space variables (such as those in the Papapetrou equations) introduces considerable complications. In particular, in unconstrained systems the deviation vector between nearby initial conditions is essentially arbitrary, but in constrained systems the deviations must be carefully constructed. The crux of the matter can be seen by considering a simplified model system: a simple pendulum described in *Cartesian coordinates* (x, y) , instead of the more natural polar coordinates $(r = \text{const.}, \theta)$.¹³ For simplicity, consider a deviation only in the spatial coordinates (i.e., not in the momenta), so that in polar coordinates we consider an initial angle θ and a nearby angle $\theta + \delta\theta$, where $\delta\theta$ is an arbitrary (small) deviation. In Cartesian coordinates, we are *not* free to consider a point $(x + \delta x, y + \delta y)$ for arbitrary deviations $(\delta x, \delta y)$; since the coordinates themselves are constrained to satisfy

$$x^2 + y^2 = r^2, \quad (1.41)$$

the deviations must satisfy

$$(x + \delta x)^2 + (y + \delta y)^2 = r^2, \quad (1.42)$$

so that $\delta y = -(x/y)\delta x$ to lowest order. (It makes sense that the components of the deviation are related, since we have added a spurious degree of freedom by using Cartesian coordinates.) In Chapter 5, we introduce the notion of a *constraint-satisfying deviation*, that is, a deviation whose components are chosen so that the deviated trajectory satisfies the same constraints as the initial point. We then show how these constraint-satisfying deviations allow all the methods for calculating Lyapunov exponents to be generalized to constrained dynamical systems.

1.2.5 The meaning of zero

Finally, we remark on the meaning of “zero” in the context of Lyapunov exponents. Chaotic systems are characterized by at least one positive Lyapunov exponent; for nonchaotic systems, all Lyapunov exponents are zero. And yet, in reality we must use prescriptions such as that in Eq. (1.36), which

¹³The simple pendulum is not chaotic, of course, but it is just complicated enough to illustrate our point.

are never exactly zero numerically, even for orbits known to be nonchaotic. The formal limit of infinite time used in Eq. (1.34) is impossible to implement in practice; the best we can do, it seems, is to calculate λ from Eq. (1.36) for large values of t , and place a bound on the size of the exponent. It turns out that we can do slightly better than this: by rearranging Eq. (1.36), we can write the growth of the deviation vector as

$$\log[r(t)] = \lambda t, \quad (1.43)$$

which means that we can calculate $\log[r(t)]$ as a function of time, and then perform a least-squares fit on the simulation data; the slope of the fit is the Lyapunov exponent. This method, since it is sensitive only to the slope and not to the absolute value, gives a better bound on λ than a naïve application of Eq. (1.36).

Another indicator of chaotic behavior is the shape of the *Lyapunov plot*, i.e., a plot of $\log[r(t)]$ vs. t . For nonzero Lyapunov exponents, such plots are roughly linear, with a slope equal to the Lyapunov exponent, as indicated by Eq. (1.43). For nonchaotic systems, such plots are logarithmic: nonchaotic separations grow as a power-law at worst, so that

$$\log[r(t)] \propto \log t^\alpha = \alpha \log t. \quad (1.44)$$

The $\log t$ vs. t behavior characterizes many of the suspected nonchaotic initial conditions discussed in this thesis.

Using the shape (as determined by visual inspection) of the $\log t$ vs. t graph to evaluate the presence of chaos is a blunt and error-prone instrument for chaos detection: even for chaotic initial conditions, the deviation initially grows as a power law, with the exponential growth dominating only at later times. It is always possible that seemingly power-law growth is chaotic on long-enough timescales. It is therefore important to calculate *baseline* orbits that are known not to be chaotic. For example, in the Papapetrou case, we calculate a baseline orbit by setting the spin parameter S to zero, corresponding to a Kerr geodesics, which is integrable and therefore must (in general) have vanishing Lyapunov exponents. The magnitude of the numerical exponents in these cases therefore gives an indication of the typical strength of power-law separation. If a suspected chaotic orbit gives an exponent with a magnitude similar to that of a nonchaotic baseline orbit, then we say that the Lyapunov exponent is “indistinguishable from zero,” and the corresponding initial condition is probably not chaotic on the timescale considered.

1.3 Chaos in general relativity: a review and critique

We present here a review of some subtleties and issues related to chaos in the context of general relativity. In the process, we review and critique some previous work on the subject, and evaluate

previous claims of chaotic behavior in spinning relativistic binaries.

1.3.1 Chaos detection

While Lyapunov exponents are our primary method for investigating chaos in relativistic binaries, there are several other possible methods. We discuss several qualitative methods (and indicate our reasons for avoiding them) in Chapter 2, but here we take the opportunity to examine the method of *fractal basin boundaries*, employed by J. Levin in a series of papers on chaos in the PN equations of motion [19, 29, 30] (the last of which was co-authored with N. J. Cornish).

In the context of spinning binaries (as used in, e.g., [19]), a basin is simply the two-dimensional space (θ_1, θ_2) , where θ is the angle between the vertical axis and each particle’s initial spin. Levin varies (θ_1, θ_2) for fixed initial values of the position and momentum, and then determines whether the resulting orbits coalesce, escape to infinity, or are stable and bounded. By color-coding each behavior and plotting points in (θ_1, θ_2) space, Levin shows that the boundary between the three types of behavior is fractal, indicating a sensitive dependence on initial conditions. For example, a fractal boundary between stable and plunge orbits means that arbitrarily small changes in the initial conditions can lead to coalescence instead of stability.

The fractal basin boundary method has several limitations. First, it makes no distinction between chaotic and nonchaotic bound orbits: a non-fractal region of the basin corresponding to bound orbits (as appears in [19]) can be full of chaos, independent of the existence of a fractal basin boundary. In other words, the key issue of the *prevalence* of chaotic orbits, so important from the perspective of gravitational-wave templates, is simply not addressed by fractal basin boundaries. Moreover, the establishment of sensitive dependence through fractal basin boundaries gives no indication of the *timescale* of the chaos—this requires a Lyapunov exponent. Finally, fractal basin boundaries only establish sensitive dependence in the asymptotic behavior of the orbits, i.e., bound, unbound, or plunge. This is completely different from the case most relevant for gravitational-wave generation: an initial condition leading to a bound orbit whose nearby initial conditions separate exponentially fast, and *also lead to bound orbits*. In fact, Cornish and Levin [31] show using Lyapunov exponents that there are such orbits located near the fractal basin boundaries identified in [19]—but this simply underscores the primacy of Lyapunov exponents as a measure of chaos.

1.3.2 Time coordinate subtleties

General relativity has a time-redefinition ambiguity, allowing any chaos apparently to be defined away by means of a coordinate transformation. If a system is chaotic when measured using time t , merely set $t' = \log t$ and the chaos seemingly disappears. Concern about this problem prevented Levin from calculating Lyapunov exponents in [19, 29]. Even in a recent paper [31], Cornish and

Levin warn of this ambiguity—and then proceed to calculate Lyapunov exponents.

Cornish and Levin are perhaps (overly) influenced by their own role in addressing chaos in mixmaster cosmological models [32], where the time ambiguity is relevant. The problem simply does not exist in our context. One can give technical arguments based on the preferred time direction picked out by the Schwarzschild or Kerr solutions (implicit also in the PN equations), but this is unnecessary: there is no ambiguity in the meaning of the Lyapunov exponents as seen by an observer at infinity. Similarly, there is no ambiguity in the meaning of the Lyapunov exponents in the local Lorentz frame of a particle, where proper time is the appropriate coordinate. Furthermore, these two descriptions are connected, since—in a static background spacetime—the coordinate time (the time-at-infinity) is simply a (spatially dependent) reparameterization of the proper time. A positive Lyapunov exponent has an unambiguous interpretation for the problem at hand.

1.3.3 Genericity and the choice of initial conditions

A couple of papers [8, 31] have shown that there are positive Lyapunov exponents in certain specialized cases for Schwarzschild *geodesics*. Such orbits are fully integrable, and hence cannot be chaotic; formal definitions of chaos explicitly rule out such nongeneric cases for precisely this reason [27]. Nevertheless, in an important sense—from the perspective of gravitational wave detection—sensitive dependence is the key property. The sensitive orbits in [8, 31], on the other hand, must be chosen with exquisite care; [8] considers *unstable, circular* geodesics, and [31] considers both that case and a more general class of unstable orbits. In both cases, it is the specialized initial conditions that lead to positive Lyapunov exponents, not a general preponderance of chaotic orbits. Such specialized initial conditions tell us nothing about the prevalence of chaos in relativistic binary systems.

Unlike the specialized unstable geodesics considered in [8, 31], the chaos found in [29] (using the PN equations) corresponded to highly eccentric comparable-mass binaries, whose relevance is potentially limited by the likely circularization of inspiraling binaries. Moreover, there is no indication in [29] of the prevalence of these chaotic initial conditions. In fact, by calculating Lyapunov exponents, Schnittman and Rasio [33] argued that there is no chaos in the PN system. Their result turned out to be too strong, as shown in [30]—there are at least some PN initial conditions with positive Lyapunov exponents. Nevertheless, the failure to find chaos in [33] is a hint that chaos is rare. This agrees with our finding (for quasi-circular orbits) presented in Chapter 4.

1.3.4 Papapetrou spin parameter confusion

The paper by Suzuki and Maeda [21], which represents the primary starting point for the project represented by Chapters 2 and 3, measures the Papapetrou spin parameter in terms of μM , as described in Sec. 1.1.3 above. They find chaos only when $S \sim 1$, but they do not emphasize the

unphysical nature of this parameter value. In fact, they consider values of S greater than 1: “Because the value of $S = 1 \mu M$ is not a mathematical special bound, we will analyze the case of $S > 1 \mu M$ as well in order to see the spin effect more clearly.” Their main concession to the limitations on their results is that, for $S = 1.4 \mu M$, “Very strong chaos occurs in this case, although it may be [physically] unrealistic.” (As shown in Chapter 2, any value of $S \gtrsim 10^{-4} \mu M$ is physically unrealistic.)

Perhaps as a result of this under-emphasis on the limitations of their results (and perhaps also due to their title, “Chaos in Schwarzschild spacetime: The motion of a spinning particle”), Ref. [21] has been cited as strong evidence of chaos in binary black hole orbits. For example, in [29] we find the following, citing [21]:

Another important example of chaos around a black hole is the motion of a spinning test particle. This already shows the key features of the two-body [post-Newtonian] system investigated here.

Other examples of overstatement can be found in [8, 31, 34]. To their credit, Cornish and Levin do note in [30] that Suzuki and Maeda find chaos only “for unphysically large values of the particle’s spin.” But one fears that the damage is done.

1.3.4.1 Chaos in this thesis

As a result of these considerations, in this thesis we use Lyapunov exponents to determine the prevalence of chaos and the corresponding timescales, we avoid the non-generic, unstable initial conditions leading to sensitive dependence in the absence of true chaos, and we indicate clearly whether the initial conditions leading to chaos are physically valid.

Bibliography

- [1] G. J. Sussman and J. Wisdom, *Science* **241**, 433 (1988).
- [2] J. Touma and J. Wisdom, *Science* **259**, 1294 (1993).
- [3] <http://www.ligo.caltech.edu>.
- [4] <http://lisa.jpl.nasa.gov>.
- [5] B. J. Owen, *Phys. Rev. D* **53**, 6749 (1996).
- [6] E. E. Flanagan and S. A. Hughes, *Phys. Rev. D* **57**, 4566 (1998).
- [7] A. Buonanno, Y. Chen, and M. Vallisneri, *Phys. Rev. D* **67**, 024016 (2003).
- [8] N. J. Cornish, *Phys. Rev. D* **64**, 084011 (2001).
- [9] T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
- [10] T. Damour and G. Schäfer, *Nuov. Cimento* **101**, 127 (1988).
- [11] T. Damour, P. Jaranowski, and G. Schäfer, *Phys. Rev. D* **62**, 084011 (2000).
- [12] E. Poisson, *Phys. Rev. D* **57**, 5287 (1998).
- [13] C. M. Will, *Living Rev. Relativity* **4** (30 Apr 2003),
<http://www.livingreviews.org/Articles/Volume4/2001-4will>.
- [14] F. D. Ryan, *Phys. Rev. D* **52**, 3159 (1995).
- [15] A. Papapetrou, *Proc. R. Soc. London* **A209**, 248 (1951).
- [16] W. G. Dixon, *Proc. R. Soc. London* **A314**, 499 (1970).
- [17] R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago, 1984).
- [18] T. Tanaka, Y. Mino, M. Sasaki, and M. Shibata, *Phys. Rev. D* **54**, 3762 (1996).
- [19] J. Levin, *Phys. Rev. Lett.* **84**, 3515 (2000).

- [20] L. E. Kidder, Phys. Rev. D **52**, 821 (1995).
- [21] S. Suzuki and K. Maeda, Phys. Rev. D **55**, 4848 (1997).
- [22] O. Semerák, Mon. Not. R. Astron. Soc. **308**, 863 (1999).
- [23] C. Møller, *Theory of Relativity* (Oxford University Press, London, 1972).
- [24] I. B. Khriplovich and A. A. Pomeransky (1996), [gr-qc/9602004](#).
- [25] B. M. Barker and R. F. O'Connell, Gen. Relativ. Gravit. **5**, 539 (1974).
- [26] K. Heinemann, [physics/9611001](#).
- [27] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems* (Springer, New York, 1997).
- [28] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1993).
- [29] J. Levin, Phys. Rev. D **67**, 044013 (2003).
- [30] N. J. Cornish and J. Levin, [gr-qc/0207016](#).
- [31] N. J. Cornish and J. Levin, [gr-qc/0304056](#).
- [32] N. Cornish and J. Levin, Phys. Rev. Lett. **78**, 998 (1997).
- [33] J. D. Schnittman and F. A. Rasio, Phys. Rev. Lett. **87**, 121101 (2001).
- [34] N. J. Cornish and J. Levin, Phys. Rev. Lett. **89**, 179001 (2002).

Chapter 2

Dynamics of spinning test particles in Kerr spacetime

This chapter was published as M. D. Hartl, Phys. Rev. D **67**, 024004 (2003); available (with corrections to the published version) as gr-qc/0210042.

Abstract

We investigate the dynamics of relativistic spinning test particles in the spacetime of a rotating black hole using the Papapetrou equations. We use the method of Lyapunov exponents to determine whether the orbits exhibit sensitive dependence on initial conditions, a signature of chaos. In the case of maximally spinning equal-mass binaries (a limiting case that violates the test-particle approximation) we find unambiguous positive Lyapunov exponents that come in pairs $\pm\lambda$, a characteristic of Hamiltonian dynamical systems. We find no evidence for nonvanishing Lyapunov exponents for physically realistic spin parameters, which suggests that chaos may not manifest itself in the gravitational radiation of extreme mass-ratio binary black-hole inspirals (as detectable, for example, by LISA, the Laser Interferometer Space Antenna).

2.1 Introduction

The presence of chaos (or lack thereof) in relativistic binary inspiral systems has received intense attention recently due to the implications for gravitational-wave detection [1–8], especially regarding the generation of theoretical templates for use in matched filters. There is concern that the sensitive dependence to initial conditions that characterizes chaos may make the calculation of such templates difficult or impossible [8]. In particular, in the presence of chaos the number of templates would increase exponentially with the number of wave cycles to be fitted. In addition to this important concern, the problem of chaos in general relativity has inherent interest, as the dynamical behavior

of general relativistic systems is poorly understood.

Several authors have reported the presence of chaos for systems of two point masses in which one or both particles are spinning [1, 3, 6]. Our work follows up on [3], which studies the dynamics of a spinning test particle orbiting a nonrotating (Schwarzschild) black hole using the Papapetrou equations [Eqs. (2.9)]. We extend this work to a rotating (Kerr) black hole, motivated by the expectation that many astrophysically relevant black holes have nonzero angular momentum. Furthermore, the potential for chaos may be greater in Kerr spacetime since the Kerr metric has less symmetry and hence fewer integrals of the motion than Schwarzschild. In addition, the decision to focus on test particles is motivated partially by the LISA gravitational wave detector [9], which will be sensitive to radiation from spinning compact objects orbiting supermassive black holes in galactic nuclei. Using the Kerr metric is appropriate since such supermassive black holes will in general have nonzero spin.

There are many techniques for investigating chaos in dynamical systems, but for the case at hand we favor the use of Lyapunov exponents to quantify chaos. Informally, if ϵ_0 is the phase-space distance between two nearby initial conditions in phase space, then for chaotic systems the separation grows exponentially (*sensitive dependence on initial conditions*): $\epsilon(\tau) = \epsilon_0 e^{\lambda\tau}$, where λ is the Lyapunov exponent. (See Sec. 2.3.1 for a discussion of issues related to the choice of metric used to determine the distance in phase space.) The value of Lyapunov exponents lies not only in establishing chaos, but also in providing a characteristic timescale $\tau_\lambda = 1/\lambda$ for the exponential separation.

By definition, chaotic orbits are bounded phase space flows with at least one nonzero Lyapunov exponent. There are additional technical requirements for chaos that rule out periodic or quasiperiodic orbits, equilibria, and other types of patterned behavior [10]. For example, unstable circular orbits in Schwarzschild spacetime can have positive Lyapunov exponents [5], but such orbits are completely integrable (see Sec. 2.6) and hence not chaotic. In practice, we restrict ourselves to generic orbits, avoiding the specialized initial conditions that lead to positive Lyapunov exponents in the absence of chaos.

The use of Lyapunov exponents is potentially dangerous in general relativity because of the freedom to redefine the time coordinate. Chaos can seemingly be removed by a coordinate transformation: simply let $\tau' = \log \tau$ and the chaos disappears. Fortunately, in our case there is a fixed background spacetime with a time coordinate that is not dynamical but rather is simply a reparameterization of the proper time. As a result, we will not encounter this time coordinate redefinition ambiguity (which plagued, for example, attempts to establish chaos in mixmaster cosmological models, until coordinate-invariant methods were developed [11]). Furthermore, we can compare times in different coordinate systems using ratios: if t_p is the period of a periodic orbit in some coordinate system with time coordinate t , and τ_p is the period in proper time, then their ratio provides a

conversion factor between times in different coordinate systems [5]:

$$\frac{t}{\tau} = \frac{t_p}{\tau_p}. \quad (2.1)$$

For chaotic orbits, which are not periodic, we use the average value of $dt/d\tau$ over the orbit, so that

$$\frac{t_\lambda}{\tau_\lambda} = \left\langle \frac{dt}{d\tau} \right\rangle \quad (2.2)$$

as discussed in Sec. 2.7.4. [This more general formula reduces to Eq. (2.1) in the case of periodic orbits.] Since we want to measure the local divergence of trajectories, the natural definition is to use the divergence in local Lorentz frames, which suggests that we use the proper time τ as our time parameter. The Lyapunov timescale in any coordinate system can then be obtained using Eq. (2.2).

Lyapunov exponents provide a quantitative definition of chaos, but there are several common qualitative methods as well, none of which we use in the present case, for reasons explained below. Perhaps the most common qualitative tool in the analysis of dynamical systems is the use of Poincaré surfaces of section. Poincaré sections reduce the phase space by one dimension by considering the intersection of the phase space trajectory with some fixed surface, typically taken to be a plane. Plotting momentum vs. position for intersections of the trajectory with this surface then gives a qualitative view of the dynamics. As noted in [4], such sections are most useful when the number of degrees of freedom minus the number of constraints (including integrals of the motion) is not greater than two, since in this case the resulting points fall on a one-dimensional curve for non-chaotic orbits, but are “dusty” for chaotic orbits (and in the case of dissipative dynamical systems lie on fractal attractors). Unfortunately, the system we consider has too many degrees of freedom for Poincaré sections to be useful. It is possible to plot momentum vs. position when the trajectory intersects a section that is a plane in physical space (say $x = 0$) [3], but this is not in general a true Poincaré section.¹

Other qualitative methods include power spectra and chaotic attractors. The power spectra for regular orbits have a finite number of discrete frequencies, whereas their chaotic counterparts are continuous. Unfortunately, it is difficult to differentiate between complicated regular orbits, quasiperiodic orbits, and chaotic orbits, so we have avoided their use. Chaotic attractors, which typically involve orbits asymptotically attracted to a fractal structure, are powerful tools for exploring chaos, but their use is limited to dissipative systems [10]. Nondissipative systems, including test particles in general relativity, do not possess attractors [12].

Following Suzuki and Maeda [3], we use the Papapetrou equations to model the dynamics

¹In [3], they are aided by the symmetry of Schwarzschild spacetime, which guarantees that one component of the spin tensor (Sec. 2.2.1 below) is zero in the equatorial plane. As a result, it turns out that all but two of their variables are determined on the surface, and thus their sections are valid. Unfortunately, the reduced symmetry of the Kerr metric makes this method unsuitable for the system we consider in this chapter.

of a spinning test particle in the absence of gravitational radiation. We extend their work in a Schwarzschild background by considering orbits in Kerr spacetime, and we also improve on their methods for calculating Lyapunov exponents. The most significant improvement is the use of a rigorous method for determining Lyapunov exponents using the linearized equations of motion for each trajectory in phase space (Sec. 2.3.1), which requires knowledge of the Jacobian matrix for the Papapetrou system (Sec. 2.5.2). We augment this method with an implementation of an informal deviation vector approach, which tracks the size of an initial deviation of size ϵ_0 and uses the relation $\epsilon(\tau) = \epsilon_0 e^{\lambda\tau}$ discussed above. We are careful in all cases to incorporate the constrained nature of the Papapetrou equations (Sec. 2.2.1) in the calculation of Lyapunov exponents (Sec. 2.4.2).

We use units where $G = c = 1$ and sign conventions as in MTW [13]. We use vector arrows for 4-vectors (e.g., \vec{p} for the 4-momentum) and boldface for Euclidean vectors (e.g., $\boldsymbol{\xi}$ for a Euclidean tangent vector). The symbol \log refers in all cases to the natural logarithm \log_e .

2.2 Spinning test particles

2.2.1 Papapetrou equations

The Papapetrou equations [14] describe the motion of a spinning test particle. Although Papapetrou first derived the equations of motion for such a particle, the formulation by Dixon [15] is the starting point for most investigations because of its conceptual clarity. Dixon writes the equations of motion in terms of the 4-momentum p^α and spin tensor $S^{\alpha\beta}$, which are defined by integrals of the particle's stress-energy tensor $T^{\alpha\beta}$ over an arbitrary spacelike hypersurface Σ :

$$p^\alpha(\Sigma) = \int_{\Sigma} T^{\alpha\beta} d\Sigma_{\beta} \quad (2.3)$$

$$S^{\alpha\beta}(\vec{z}, \Sigma) = 2 \int_{\Sigma} (x^{[\alpha} - z^{[\alpha}) T^{\beta]\gamma} d\Sigma_{\gamma}, \quad (2.4)$$

where \vec{z} is the coordinate of the center of mass. The equations of motion for a spinning test particle are then

$$\begin{aligned} \frac{dx^\mu}{d\tau} &= v^\mu \\ \nabla_{\vec{v}} p^\mu &= -\frac{1}{2} R^\mu{}_{\nu\alpha\beta} v^\nu S^{\alpha\beta} \\ \nabla_{\vec{v}} S^{\mu\nu} &= 2p^{[\mu} v^{\nu]}, \end{aligned} \quad (2.5)$$

where v^μ is the 4-velocity, i.e., the tangent to the particle's worldline. It is apparent that the 4-momentum deviates from geodesic motion due to a coupling of the spin to the Riemann curvature.

2.2.1.1 Spin supplementary conditions

As written, the Papapetrou equations 2.5 are under-determined, and require a *spin supplementary condition* to determine the rest frame of the particle's center of mass. Following Dixon, we choose

$$p_\mu S^{\mu\nu} = 0, \quad (2.6)$$

which picks out a unique worldline that we identify as the center of mass. In particular, in the zero 3-momentum frame defined by $p^i = 0$, applying Eq. (2.6) to Eq. (2.4) yields

$$z^i = \frac{\int_{t=\text{const.}} x^i T^{00} d^3x}{\int_{t=\text{const.}} T^{00} d^3x}, \quad (2.7)$$

which is the proper relativistic generalization of the Newtonian center of mass. The frame defined by $p^i = 0$ is thus the rest frame of the center of mass, and in this frame Eq. (2.6) implies that $S^{0j} = 0$, i.e., the spin is purely spatial in the rest frame.

A second possibility for the supplementary condition is

$$v_\mu S^{\mu\nu} = 0. \quad (2.8)$$

This condition has the disadvantage that it is satisfied by a family of helical worldlines filling a cylinder with frame-dependent radius [15, 16], centered on the worldline picked out by condition 2.6. As a result, we adopt $p_\mu S^{\mu\nu} = 0$ as the supplementary condition.

We note that the difference between the conditions 2.6 and 2.8 is third order in the spin [which follows from Eq. (2.15) below], which means that it is negligible for physically realistic spins (Sec. 2.2.2). In particular, the two conditions are equivalent for post-Newtonian expansions [17], where condition 2.8 is typically employed [18].

2.2.1.2 A reformulation of the equations

For numerical reasons, we use a form of the equations different from Eqs. (2.5). (We discuss this and other numerical considerations in Sec. 2.5.1.) Following the appendix in [3], we write the equations in terms of the momentum 1-form p_μ and the spin 1-form S_μ .² The system under consideration is a spinning particle of rest mass μ orbiting a central body of mass M ; in what follows, we measure all times and lengths in terms of M , and we measure the momentum of the particle in terms of μ ,

²The lowered indices are motivated by the Hamiltonian formulation for a nonspinning test particle, where it is the one-form p_μ that is canonically conjugate to x^μ [13].

so that $p_\nu p^\nu = -1$. In these normalized units, the equations of motion are

$$\begin{aligned}\frac{dx^\mu}{d\tau} &= v^\mu \\ \nabla_{\vec{v}} p_\mu &= -R_{\mu\nu}^{*\alpha\beta} v^\nu p_\alpha S_\beta \\ \nabla_{\vec{v}} S_\mu &= -p_\mu \left(R_{\beta}^{*\alpha\gamma\delta} S_\alpha v^\beta p_\gamma S_\delta \right),\end{aligned}\tag{2.9}$$

where

$$R_{\beta}^{*\alpha\mu\nu} = \frac{1}{2} R_{\beta\rho\sigma}^{\alpha} \epsilon^{\rho\sigma\mu\nu}.\tag{2.10}$$

The tensor and vector formulations of the spin are related by

$$S_\mu = -\frac{1}{2} \epsilon_{\mu\nu\alpha\beta} u^\nu S^{\alpha\beta}\tag{2.11}$$

and

$$S^{\mu\nu} = -\epsilon^{\mu\nu\alpha\beta} S_\alpha u_\beta,\tag{2.12}$$

where $u_\nu = p_\nu/\mu$ ($= p_\nu$ in normalized units). In addition, the spin satisfies the condition

$$S_\mu S^\mu = \frac{1}{2} S_{\mu\nu} S^{\mu\nu} = S^2,\tag{2.13}$$

where S is the spin of the particle measured in units of μM (see Sec. 2.2.2).

Because of the coupling of the spin to the Riemann curvature, the 4-momentum p^μ [Eq. (2.3)] is not parallel to the tangent v^μ . The supplementary condition 2.6 allows for an explicit solution for the difference between them (see [19] for a derivation):

$$v^\mu = N(p^\mu + w^\mu),\tag{2.14}$$

where

$$w^\mu = -{}^*R^{*\mu\alpha\beta\gamma} S_\alpha p_\beta S_\gamma\tag{2.15}$$

and

$${}^*R^{*\alpha\beta\mu\nu} = \frac{1}{2} R^{*\alpha\beta\rho\sigma} \epsilon_{\rho\sigma}{}^{\mu\nu}.\tag{2.16}$$

The normalization constant N is fixed by the constraint $v_\mu v^\mu = -1$. We see from Eq. (2.15) that the difference between p^μ and v^μ is $\mathcal{O}(S^2)$, so that the difference between Eqs. (2.6) and 2.8 is $\mathcal{O}(S^3)$.

The spin 1-form satisfies two orthogonality constraints:

$$p^\mu S_\mu = 0\tag{2.17}$$

and

$$v^\mu S_\mu = 0. \quad (2.18)$$

These two constraints are equivalent as long as v^μ is given by Eq. (2.14), since $w^\mu S_\mu \propto {}^*R^{\mu\alpha\beta\gamma} S_\mu S_\alpha \equiv 0$. When parameterizing the initial conditions, we enforce Eq. (2.17); since we use Eq. (2.14) in the equations of motion, Eq. (2.18) is then automatically satisfied.

2.2.1.3 Range of validity

We note that the Papapetrou equations include effects due only to the mass monopole and spin dipole (the pole-dipole approximation). In particular, the tidal coupling, which is a mass quadrupole effect, is neglected. It is also important to note that the Papapetrou equations are conservative and hence ignore the effects of gravitational radiation. For a thorough and accessible general discussion of the Papapetrou equations and related matters, including a comprehensive literature review, see Semerák [19].

2.2.2 Comments on the spin parameter

It is crucial to note that, in our normalized units, the spin parameter S is measured in terms of μM , not μ^2 . The system we consider in this chapter is a compact spinning body of mass μ orbiting a large body of mass M , which we take to be a supermassive Kerr black hole satisfying $M \approx 10^5\text{--}10^6 M_\odot$. We will show that physically realistic values of the spin must satisfy $S \ll 1$ for the compact objects (black holes, neutron stars, and white dwarfs) most relevant for the test particles described by the Papapetrou equations.³ The case of a black hole is simplest: a maximally spinning black hole of mass μ has spin angular momentum $s = \mu^2$, so a small black hole μ orbiting a large black hole of mass $M \gg \mu$ has a small spin parameter S :

$$S = \frac{s}{\mu M} \leq \frac{\mu^2}{\mu M} = \frac{\mu}{M} \ll 1.$$

The limit is similar for neutron stars: most models of neutron stars have a maximum spin of $s_{\max} \approx 0.6 \mu^2$ [20], which gives $S \lesssim 0.6 \mu/M$.

2.2.2.1 Bounds on S for stellar objects

The bound on S is relatively simple for black holes and neutron stars, but the situation is more complicated for compact stellar objects such as white dwarfs. The maximum spin of a stellar object is typically determined by the mass-shedding limit, i.e., the maximum spin before the star begins to break up. The spin in the case of the break-up limit is the moment of inertia times the maximum

³Recall that the Papapetrou equations ignore tidal coupling, so they are inappropriate for modeling more extended objects.

(break-up) angular velocity: $s_{\max} = I\Omega_{\max}$. If we write $I = \alpha\mu R^2$ and $\Omega_{\max} = \beta\sqrt{G\mu/R^3}$ for some constants $\alpha, \beta \lesssim 1$, then we have

$$s_{\max} = \alpha\beta (G\mu^3 R)^{1/2}. \quad (2.19)$$

The values of α and β depend on the stellar model; if we use the values for an $n = 1.5$ polytrope, we get $\alpha = 0.2044$ and $\beta = 0.5366$ [21], so that $s_{\max} = 0.110 (G\mu^3 R)^{1/2}$.

The limit in Eq. (2.19) depends on the mass-radius relation for the object in question. Since most neutron stars have masses and radii in a narrow range, the estimate of $s_{\max} \approx 0.6\mu^2$ discussed above is sufficient, but for white dwarfs the value of s_{\max} can depend strongly on the mass. An analytical approximation for the mass-radius relation for non-rotating white dwarfs is [22]:⁴

$$\frac{R}{R_{\odot}} = 0.01125 \left(\frac{\mu}{\mu_{\max}} \right)^{-1/3} f(\mu)^{1/2} \quad (2.20)$$

where

$$f(\mu) = 1 - \left(\frac{\mu}{\mu_{\max}} \right)^{4/3} \quad (2.21)$$

and

$$\mu_{\max} = 1.454 M_{\odot}. \quad (2.22)$$

We could plug Eq. (2.20) into Eq. (2.19) to obtain an order-of-magnitude estimate, but [21] tabulates a constant \bar{J} equivalent to the product $\alpha\beta$ (which increases as the angular velocity of the star increases). They write $J = \bar{J} (GM^3 R_0)^{1/2}$ for a rotating white dwarf, where \bar{J} depends on the polytropic index n of a *non-spinning* white dwarf of the same mass, and R_0 is the non-spinning radius. In our notation, this reads

$$s_{\max} = \bar{J} (G\mu^3 R)^{1/2}. \quad (2.23)$$

White dwarfs with $\mu > 0.6 M_{\odot}$ are not well approximated by polytropes (the effective polytropic index varies from near 3 in the core to near 1.5 in the outer parts), but useful bounds can be obtained by substituting R from Eq. (2.20), which is more accurate for white dwarfs than a pure polytrope model. Plugging 2.20 into 2.23 and converting to geometric units gives

$$s_{\max} = 77.68 \bar{J} \mu^{4/3} M_{\odot}^{2/3} f(\mu)^{1/4}. \quad (2.24)$$

From Table 3 in [21], we have $\bar{J} = 0.1660$ for a maximally rotating $n = 1.5$ polytrope (vs. $\alpha\beta = 0.110$ for a slowly rotating one) and $\bar{J} = 0.0785$ for $n = 2.5$. As illustrated in Fig. 2.1, the values for a more realistic numerical model [23] lie between these curves, as expected.

Note from Eq. (2.24) that $s_{\max}/\mu^2 \propto \mu^{-2/3}$ for $\mu \ll \mu_{\max}$, so that the spin per unit mass squared

⁴The mean molecular weight $\bar{\mu}$ is set equal to 2, corresponding to helium and heavier elements, which is appropriate for most astrophysical white dwarfs.

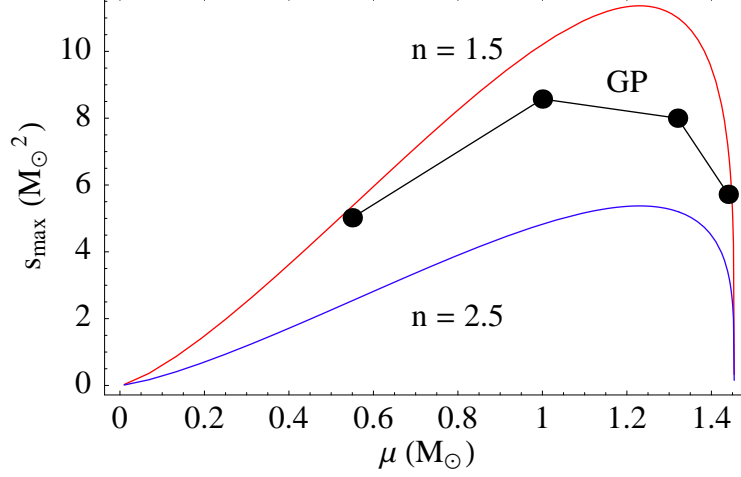


Figure 2.1: The maximum spin angular momentum s_{\max} vs. mass μ for a rigidly rotating white dwarf. We plot curves for $n = 1.5$ and $n = 2.5$ polytropic approximations using Eq. (2.24), together with four points derived using a more realistic numerical white dwarf model (Geroyannis and Papa-sotiriou [23]).

is unbounded as $\mu \rightarrow 0$.⁵ Nevertheless, the spin parameter S_{\max} is bounded, since $S_{\max} \propto s_{\max}/\mu \propto \mu^{1/3}$ in the low mass limit. We plot s_{\max}/μ vs. μ in Fig. 2.2, which shows that the maximum value of s_{\max}/μ is approximately $9 M_{\odot}$ (corresponding to a $\mu = 0.5 M_{\odot}$ white dwarf). For a central black hole of mass $M = 10^6 M_{\odot}$, we then have

$$S \leq S_{\max} = \frac{s_{\max}}{\mu M} = 9 \times 10^{-6}, \quad (2.25)$$

which is small compared to unity.

2.2.2.2 Tidal disruption

We can obtain a higher value of S if the central black hole mass is smaller, but it is important to bear in mind that such lower-mass black holes may tidally disrupt the white dwarf companion, thereby violating a necessary condition for the validity of the Papapetrou equations. In order of magnitude, a white dwarf orbiting at radius r will be disrupted when the tidal acceleration due to the central body overcomes its self-gravity, i.e.,

$$\frac{GM}{r^3} R \geq \frac{G\mu}{R^2}. \quad (2.26)$$

For the white dwarf to be undisrupted down to the horizon at $r = M$, we must have $M \leq R^{3/2} \mu^{-1/2}$, so that [using Eq. (2.20)] the minimum mass not to disrupt is $M_{\min} \propto \mu^{-1}$. We could evaluate the proportionality constant using Eq. (2.20), but we can obtain a more accurate result by adopting a

⁵Eq. (2.24) is valid only for $\mu \geq 0.01 M_{\odot}$, but s_{\max}/μ^2 continues to increase with decreasing μ for equations of state appropriate for brown dwarfs and planets.

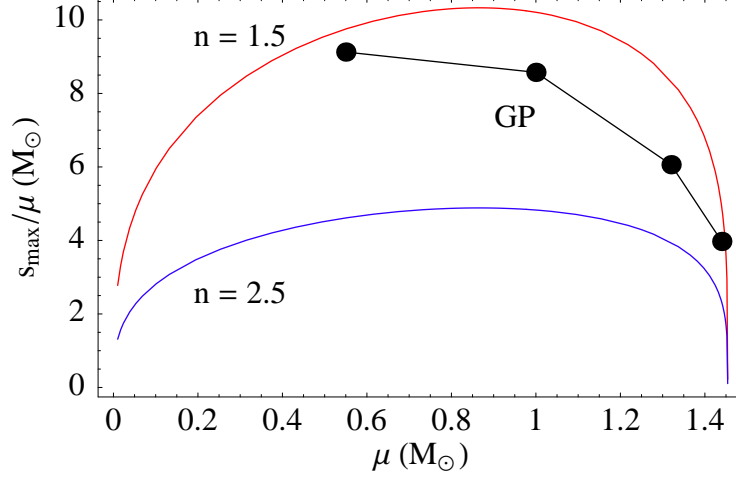


Figure 2.2: s_{\max}/μ vs. mass for a white dwarf. As in Fig. 2.1, we plot curves for $n = 1.5$ and $n = 2.5$ polytropes and the numerical model from [23]. The corresponding spin parameter S_{\max} is obtained simply by dividing s_{\max}/μ by the mass M of the central black hole.

constant based on a more realistic tidal disruption model. Tables 1 and 2 of [24] give the value of the variable $\hat{r} \equiv \frac{r}{R} \left(\frac{\mu}{M} \right)^{1/3}$, which is approximately 2.0 for the white dwarfs of interest here. This gives

$$M_{\min} = 2.0^{-3/2} R^{3/2} \mu^{-1/2}, \quad (2.27)$$

as illustrated in Fig. 2.3. For a $1.0 M_{\odot}$ white dwarf, which (based on [23]) has $s_{\max} = 8.57 M_{\odot}^2$, the central black hole must satisfy $M_{\min} = 8.2 \times 10^4 M_{\odot}$, so that the spin parameter S can be no bigger than $S_{\max} = s_{\max}/(\mu M_{\min}) = 1.0 \times 10^{-4}$ in order to avoid tidal disruption.

2.2.2.3 The $S = 1$ limit

We have shown that all physically realistic cases satisfy $S \ll 1$, but we nevertheless consider the limit of $S = 1$ (corresponding to $\mu = M$) in order to investigate more thoroughly the dynamics of the Papapetrou equations, and to compare our results with [3], which investigates the $S = 1$ limit in detail. The $S = 1$ limit introduces no singularities into the equations of motion, and the resulting orbits are valid solutions of the equations. On the other hand, in this limit the Papapetrou equations are not physically realistic, since they are derived in the limit of spinning test particles, which must satisfy $\mu \ll M$. We thus cannot draw reliable results about the behavior of astrophysical systems from the $S = 1$ limit.

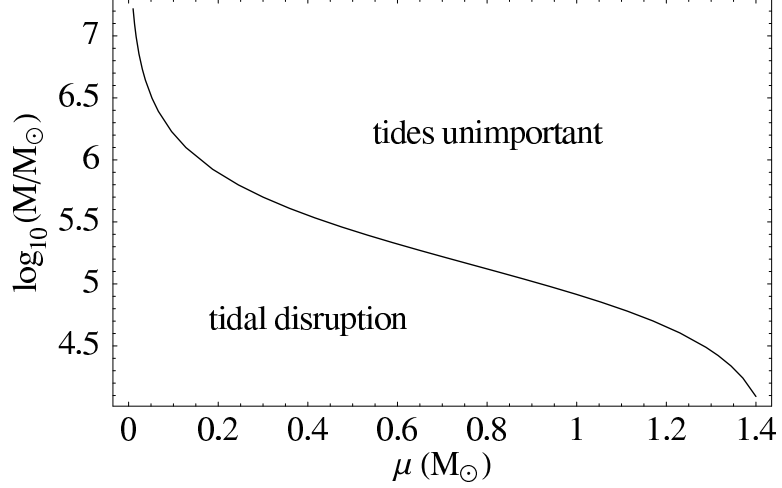


Figure 2.3: The minimum black hole mass M required not to disrupt an inspiraling corotating white dwarf before the last stable (prograde) circular orbit around a maximally rotating Kerr black hole, as a function of white dwarf mass μ .

2.2.3 Symmetries and the parameterization of initial conditions

In the approximation represented by the Papapetrou equations there is still a constant of the motion associated with each Killing vector $\vec{\xi}$ of the spacetime [15]:

$$C_{\xi} = \xi^{\mu} p_{\mu} - \frac{1}{2} \xi_{\mu;\nu} S^{\mu\nu}. \quad (2.28)$$

[For brevity, we write the constant in terms of the spin tensor $S^{\mu\nu}$ [Eq. (2.12)].] Since Kerr spacetime is stationary and axially symmetric, it has the Killing vectors $\vec{\xi}^t = \partial/\partial t$ and $\vec{\xi}^{\phi} = \partial/\partial \phi$, so the energy E and z angular momentum J_z are conserved:

$$E = -p_t + \frac{1}{2} g_{t\mu,\nu} S^{\mu\nu} \quad (2.29)$$

and

$$J_z = p_{\phi} - \frac{1}{2} g_{\phi\mu,\nu} S^{\mu\nu}. \quad (2.30)$$

(We write J_z in place of the orbital angular momentum L_z since the spin also contributes to the angular momentum of the system.) In contrast to the energy and momentum integrals, the Carter constant Q is no longer present when the test particle has nonvanishing spin [25].

In our problem there are twelve variables, four each for position, momentum, and spin. For the purposes of finding orbits by numerical integration, we may parameterize the initial conditions by providing $\tau = 0$, r , θ , $\phi = 0$, p_r , E , J_z , S , and any two of the spin components. The normalization conditions $p_{\mu} p^{\mu} = -1$ and $S_{\mu} S^{\mu} = S^2$ allow us to eliminate one component each of momentum and

spin. The constraint $p^\mu S_\mu = 0$ and the integrals of the motion then give three equations in three unknowns:

$$0 = p_\mu S_\nu g^{\mu\nu} \quad (2.31)$$

$$E = -p_t + \frac{1}{2} g_{t\mu,\nu} S^{\mu\nu} \quad (2.32)$$

$$J_z = p_\phi - \frac{1}{2} g_{\phi\mu,\nu} S^{\mu\nu} \quad (2.33)$$

We must solve these equations for the two remaining components of p_μ and one remaining component of S_μ . In Schwarzschild spacetime these can be solved explicitly due to the greater symmetry [3], but in the Kerr case of interest here the problem requires numerical root finding.

We also use a related parameterization method starting with the Kerr geodesic orbital parameters: eccentricity e , inclination angle ι , and pericenter r_p . We derive the corresponding energy, angular momentum, and relevant momenta, and then proceed as above. This method is discussed further in Sec. 2.7.1.3.

2.3 Lyapunov exponents

2.3.1 General discussion of Lyapunov exponents

Our method for calculating Lyapunov exponents is well established in the literature of nonlinear dynamical systems [10, 12], but accessible treatments are hard to find in the physics literature, so we summarize the method here. Our discussion is informal and oriented toward practical calculation, based on Ref. [10]; for a more formal, rigorous presentation see Eckmann and Ruelle [26].

First we give an overview of the methods for calculating Lyapunov exponents most commonly used in physics. Given an initial condition, a set of differential equations determines a solution (the *flow*), which is a curve in the phase space. The *Lyapunov exponents* of the flow measure the rate at which nearby trajectories separate. As discussed in the introduction, an orbit is chaotic if a nearby phase-space trajectory separated by an initial distance ϵ_0 separates exponentially: $\epsilon(\tau) = \epsilon_0 e^{\lambda\tau}$, where λ is the Lyapunov exponent.

Implicit in the definition of chaos above is a notion of a distance function on the phase space (or, more properly, the tangent space to the phase space, as in Eq. (2.36) below). It is conventional to use a Euclidean metric to define such lengths [10, 12], but any positive-definite nondegenerate metric will do [26]. While the magnitude of the resulting exponent obviously depends on the particular metric used, the signs of the Lyapunov exponents are a property of the dynamical system and do not rely on any underlying metric structure. We discuss these issues further in Sec. 2.4.1 and Sec. 2.7.4.

This informal definition of Lyapunov exponents leads to a practical method for calculating λ : given an initial condition, consider a nearby initial condition a distance ϵ_0 away, where ϵ_0 is “small”,

typically 10^{-5} – 10^{-7} of the relevant physical scales. (Values of ϵ_0 much smaller than this can result in a loss of numerical precision.) Keeping track of the deviation vector between the two points yields a numerical approximation of λ . (It is important to rescale the deviation vector if it grows too large, since for any bounded phase space flow even a tiny deviation can grow to at most the size of the bounded region.) We call this approach the *deviation vector method*.

There are two primary limitations to the approach outlined above. First, the method yields only the largest Lyapunov exponent, which is sufficient to establish the presence of chaos but paints a limited picture of the dynamics. Second, the deviation vector approach is most appropriate when an analytical expression for the Jacobian matrix is unknown; by choosing ϵ_0 small enough (and by keeping $\epsilon(\tau)$ small by rescaling if necessary), the method essentially takes a numerical derivative. Among other complications, the value of the exponent depends both on the maximum allowable size ϵ_{\max} (the size at which the deviation is rescaled) and the initial value ϵ_0 (the size of the deviation after each rescaling).

The principal virtue of the deviation vector approach compared to the more complicated Jacobian method (discussed below) is speed, since it requires solving only the equations of motion. (As we discuss in Sec. 2.3.2.1, the Jacobian method involves the time-consuming evolution of the Jacobian matrix in parallel with the equations of motion.) It also provides a valuable way to verify the validity of the Jacobian method.

The *Jacobian method* is a more thorough and rigorous approach to the calculation of Lyapunov exponents, which makes precise the notion of “infinitesimally” separated vectors. The general method proceeds as follows: consider a phase space with variables $\mathbf{y} = \{y_i\}$ and an autonomous set of differential equations

$$\frac{d\mathbf{y}}{d\tau} = \mathbf{f}(\mathbf{y}). \quad (2.34)$$

(Here we use τ instead of t in anticipation of the application of these results to general relativity, where we will be using proper time as our time parameter.) If $\delta\mathbf{y}$ represents a small deviation vector, then the distance between the two trajectories is

$$\frac{d(\delta\mathbf{y})}{d\tau} = \mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) = \mathbf{Df} \cdot \delta\mathbf{y} + \mathcal{O}(\|\delta\mathbf{y}\|^2), \quad (2.35)$$

where \mathbf{Df} is the Jacobian matrix $[(\mathbf{Df})_{ij} = \partial f_i / \partial x^j]$.

We can clarify the notation and make the system easier to visualize if we introduce $\boldsymbol{\xi}$ as an element of the tangent space at \mathbf{y} , so that

$$\frac{d\boldsymbol{\xi}}{d\tau} = \mathbf{Df} \cdot \boldsymbol{\xi}, \quad (2.36)$$

which is equivalent to taking the limit $\|\delta\mathbf{y}\| \rightarrow 0$. We visualize $\boldsymbol{\xi}$ as a perfectly finite vector (as

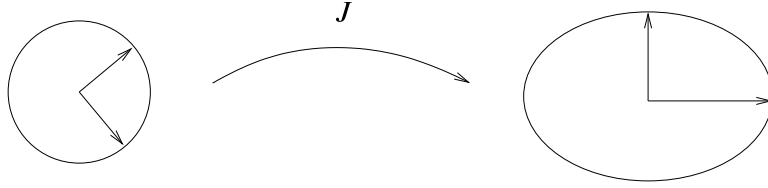


Figure 2.4: The Jacobian matrix maps a tangent space ball to an ellipsoid.

opposed to an “infinitesimal”). Since it lives in the tangent space, not the physical phase space, ξ can grow arbitrarily large with time. This means that instead of the frequent rescaling required in the deviation vector approach, ξ must be rescaled only when it grows so large that it approaches the floating point limit of the computer. This is a rare occurrence, and in practice the tangent vector almost never needs rescaling.

Although following the evolution of an arbitrary initial tangent vector ξ yields the largest Lyapunov exponent, we can do even better by following the evolution of a family of n tangent vectors, which allows us to determine all n Lyapunov exponents. The essence of the method is as follows: for a system of differential equations with n variables, we consider a set of n vectors that lie on a ball in the tangent space. We represent this ball using a matrix whose columns are n normalized, linearly independent tangent vectors, conventionally taken to be orthogonal. This set of orthonormal vectors then spans a unit ball in the tangent space. The action of the Jacobian matrix, which is a linear operator on the tangent space, is to map the ball to an ellipsoid under the time-evolution of the flow, as shown in Fig. 2.4.

For a dynamical system with n degrees of freedom, there are n *Lyapunov numbers* that measure the average growth of the n principal axes $\{r_i(\tau)\}_{i=1}^n$ of the ellipsoid. More formally, the Lyapunov numbers L_i are given by

$$L_i = \lim_{\tau \rightarrow \infty} [r_i(\tau)]^{1/\tau}, \quad (2.37)$$

where $r_i(\tau)$ is the length of the i th principal axis of the ellipsoid. The corresponding Lyapunov exponents are the natural logarithms of the Lyapunov numbers, so that

$$\lambda_i = \lim_{\tau \rightarrow \infty} \frac{\log [r_i(\tau)]}{\tau}. \quad (2.38)$$

These limits exist for a broad class of dynamical systems [26].

The principal axes of the tangent space ellipsoid indicate the directions along which nearby initial conditions separate or converge, which we may call the *Lyapunov directions*. In particular, consider a principal axis that is stretched under the time evolution. Such a vector has one component for each dimension (position or momentum) in the phase space; a nonzero component in any direction indicates an exponential divergence in the corresponding coordinate. For example, if a system has

two spatial coordinates (r, ϕ) and corresponding momenta (p_r, p_ϕ) , then a typical tangent vector will have components $\boldsymbol{\xi} = (\xi_r, \xi_\phi, \xi_{p_r}, \xi_{p_\phi})$. If the only tangent vector with nonzero Lyapunov exponent is, for example, $\boldsymbol{\xi} = \frac{1}{\sqrt{3}}(1, 1, 1, 0)$, then nearby initial conditions separate exponentially in r , ϕ , and p_r , but nearby values of p_ϕ do not separate exponentially. This is potentially relevant to the present study since, in the limit of a point test particle, the gravitational radiation depends on the spatial variables but not the spin. If the principal axes along expanding directions have nonzero components only in the spin directions, the system could be formally chaotic without affecting the gravitational waves.

In summary, the method for visualizing the Lyapunov exponents of a dynamical system is to picture a ball of initial conditions—an infinitesimal ball if visualized in the phase space, or a unit ball if visualized in the tangent space—and watch it evolve into an ellipsoid under the action of the Jacobian matrix. After a sufficiently long time, the ellipsoid will be greatly deformed, stretched out along the expanding directions and compressed along the contracting directions. The directions of the principal axes are the Lyapunov directions, and their lengths give the Lyapunov numbers through the relation $L_i \approx [r_i(\tau)]^{1/\tau}$.

2.3.2 Numerical calculation of Lyapunov exponents

In order to implement a numerical algorithm based on the considerations above, we must bear two things in mind. First, since the vectors spanning the initial unit ball are arbitrary, they will all be stretched in the direction of the largest exponent: in general every initial vector has some nonzero component along the direction of greatest stretching, which dominates as $\tau \rightarrow \infty$. In order to find the other principal axes, we must periodically produce a new orthogonal basis. We will show that the Gram-Schmidt procedure is appropriate. Second, the lengths of the vectors could potentially overflow or underflow the machine precision, so we should periodically normalize the ellipsoid axes.

2.3.2.1 The algorithm in detail

To simplify the notation, we denote the (time-dependent) Jacobian matrix \mathbf{Df} by J_τ and the ellipsoid (whose columns are the tangent vectors) by U . The algorithm then proceeds as follows:

1. Construct a set of n orthonormal vectors (which span an n -dimensional ball in the tangent space of the flow). Represent this ball by a matrix U whose columns are the tangent vectors $\boldsymbol{\xi}_i$.
2. Eq. (2.36), applied to each tangent vector, implies that U satisfies the matrix equation

$$\frac{dU}{d\tau} = J_\tau U, \quad (2.39)$$

which constitutes a set of *linear* differential equations for the tangent vectors. Since J_τ depends

on the values of \mathbf{y} , these equations are coupled to our system of nonlinear differential equations $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$, so they must be solved in parallel with Eq. (2.34).

3. Choose some time T big enough to allow the expanding directions to grow but small enough so that they are not too big. Numerically integrate Eqs. (2.34) and (2.39), and every time T apply the Gram-Schmidt orthogonalization procedure. The vectors resulting from the Gram-Schmidt procedure approximate the semiaxes of the evolving ellipsoid. Record the log of the length $\log[r_i(\tau_n)]$ of each vector after each time T , where $\tau_n = nT$. Finally, normalize the ellipsoid back to a unit ball.

4. At each time τ , the sum

$$\lambda_i \approx \frac{1}{\tau} \sum_{n=1}^N \log[r_i(\tau_n)] \equiv \frac{\log[r_i(\tau)]}{\tau} \quad (2.40)$$

is a numerical estimate for the i th Lyapunov exponent.

2.3.2.2 Gram-Schmidt and Lyapunov exponents

The use of the Gram-Schmidt procedure is crucial to extracting all n Lyapunov exponents. Let us briefly review this important construction. Given n linearly independent vectors $\{\mathbf{u}_i\}$, the Gram-Schmidt procedure constructs n orthogonal vectors $\{\mathbf{v}_i\}$ that span the same space, given by

$$\mathbf{v}_i = \mathbf{u}_i - \sum_{j=1}^{i-1} \frac{\mathbf{u}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_j\|^2} \mathbf{v}_j. \quad (2.41)$$

To construct the i th orthogonal vector, we take the i th vector from the original set and subtract off its projections onto the previous $i - 1$ vectors produced by the procedure.

The use of Gram-Schmidt in dynamics comes from observing that the resulting vectors approximate the semiaxes of the tangent space ellipsoid. After the first time T , all of the vectors point mostly along the principal expanding direction. We may therefore pick any one as the first vector in the Gram-Schmidt algorithm, so choose $\boldsymbol{\xi}_1 \equiv \mathbf{u}_1$ without loss of generality. If we let \mathbf{e}_i denote unit vectors along the principal axes and let r_i be the lengths of those axes, the dynamics of the system guarantee that the first vector \mathbf{u}_1 satisfies

$$\mathbf{u}_1 = r_1 \mathbf{e}_1 + r_2 \mathbf{e}_2 + \cdots \approx r_1 \mathbf{e}_1 \equiv \mathbf{v}_1$$

since \mathbf{e}_1 is the direction of fastest stretching. The second vector \mathbf{v}_2 given by Gram-Schmidt is then

$$\mathbf{v}_2 = \mathbf{u}_1 - \frac{\mathbf{u}_1 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 \approx \mathbf{u}_1 - r_1 \mathbf{e}_1 = r_2 \mathbf{e}_2,$$

with an error of order r_2/r_1 . The procedure proceeds iteratively, with each successive Gram-Schmidt

step (approximately) subtracting off the contribution due to the previous semiaxis direction.

It is important to choose time T long enough to keep errors of the form r_2/r_1 small but short enough to prevent numerical under- or overflow. In practice, the method is quite robust, and it is easy to find valid choices for the time T , as discussed in Sec. 2.7.

2.4 Relativity and Papapetrou subtleties

The algorithm described above is of a general nature, designed with a generic dynamical system in mind. The Papapetrou equations and the framework of general relativity present additional complications. Here we discuss some refinements to the algorithm necessary for the present case.

2.4.1 Phase space norm

In the context of general relativistic dynamical systems, the meaning of trajectory separation in phase space is somewhat obscured by the time variable. We can skirt the issue of trajectories “diverging in time” by using a $3 + 1$ splitting of spacetime, and consider trajectory separation in a spacelike hypersurface [27]. This prescription reduces properly to the traditional method for classical dynamical systems in the nonrelativistic limit.

In Kerr spacetime, we use the zero angular-momentum observers (ZAMOs), and project 4-dimensional quantities into the ZAMO hypersurface using the projection tensor $P^\mu_\nu = \delta^\mu_\nu + U^\mu U_\nu$, where U^μ is the ZAMO 4-velocity. In this formulation, spatial variables obey $x^\mu \rightarrow \tilde{x}^i = P^i_\mu x^\mu$ and momenta obey $p_\mu \rightarrow \tilde{p}_i = P^\mu_i p_\mu$ (and similarly for S_μ) [27]. The relevant norm is then a Euclidean distance in the three-dimensional hypersurface.

We should note that we use the projected norm for conceptual clarity, not necessity. The naïve use of a Euclidean norm using unprojected components yields the same sign for the exponents, as noted in Sec. 2.3.1. The magnitudes of the resulting exponents are also similar (Sec. 2.7.4).

2.4.2 Constraint complications

Although the Lyapunov algorithm is fairly straightforward to implement for a general dynamical system, the constrained nature of the Papapetrou equations adds a considerable amount of complexity. The fundamental problem is that the tangent vector ξ cannot have arbitrary initial components for the Papapetrou system, as it can for an unconstrained dynamical system. Each ξ must correspond to some deviation $\delta\mathbf{y}$ which is not arbitrary: the deviated point $\mathbf{y} + \delta\mathbf{y}$ must satisfy the constraints.

2.4.2.1 Constraint-satisfying deviations

Recall that the dynamical variables in the Papapetrou equations must satisfy normalization and orthogonality constraints (Sec. 2.2.1): $p^\nu p_\nu = -1$ (normalized units), $S^\nu S_\nu = S^2$, and $p^\nu S_\nu = 0$. To make the constraint condition on $\delta\mathbf{y}$ clearer, let $\mathbf{C}(\mathbf{y}) = 0$ represent the constraints rearranged so that the right hand side is zero. For example, with $\mathbf{y} = (t, r, \mu, \phi, p_t, p_r, p_\mu, p_\phi, S_t, S_r, S_\mu, S_\phi)$,⁶ we can write

$$C_1(\mathbf{y}) = p^\nu p_\nu + 1, \quad (2.42)$$

so that $C_1(\mathbf{y}) = 0$ for a constraint-satisfying \mathbf{y} . The other constraints are then

$$C_2(\mathbf{y}) = S^\nu S_\nu - S^2 \quad (2.43)$$

and

$$C_3(\mathbf{y}) = p^\nu S_\nu. \quad (2.44)$$

A deviation $\delta\mathbf{y}$ is *constraint-satisfying* if $\mathbf{C}(\mathbf{y} + \delta\mathbf{y}) = \mathbf{0}$ when $\mathbf{C}(\mathbf{y}) = \mathbf{0}$.

We may construct a constraint-satisfying deviation $\delta\mathbf{y}$ as follows. Begin with a 12-dimensional vector \mathbf{y} that satisfies the constraints. Add a random small deviation to eight of its components to form a new vector \mathbf{y}' . (We need not add a deviation to t ; see Sec. 2.4.2.2 below.) Determine the remaining three components of \mathbf{y}' using the constraints, using the same technique used to set the initial conditions. Finally, set $\delta\mathbf{y} \equiv \mathbf{y}' - \mathbf{y}$. The corresponding $\boldsymbol{\xi}$ is then simply $\delta\mathbf{y}/\|\delta\mathbf{y}\|$.

The prescription above glosses over an important detail: the inference of tangent vector components from the constraints is not unique. Solving the constraint equations involves taking square roots in several places, so there are a number of sign ambiguities representing different solution branches. The implementation of the component-inference algorithm must compare each component of \mathbf{y} with the corresponding component of \mathbf{y}' to ensure that they represent solutions from the same branches. Enforcing the constraints in this manner, and thereby inferring the full tangent vector $\boldsymbol{\xi}$, is especially important for the algorithm described in the next section.

2.4.2.2 A modified Gram-Schmidt algorithm

A spinning test particle has an apparent twelve degrees of freedom—four each for position, momentum, and spin—so *a priori* there is the potential for twelve nonzero exponents. Since the Papapetrou equations have no explicit time-dependence, we can eliminate the time degree of freedom. The three constraints (momentum and spin normalization, and momentum-spin orthogonality) further reduce the number of degrees of freedom by three. We are left finally with eight degrees of freedom.

Eliminating the four spurious degrees of freedom from the tangent vectors presents a formidable

⁶Recall that we write the equations of motion in terms of $\mu = \cos \theta$.

obstacle to the implementation of the phase space ellipsoid method described in Sec. 2.3.2.1. The crux of the dilemma is that the axes of the ellipsoid must be orthogonal, but must also correspond to constraint-satisfying deviation vectors—mutually exclusive conditions. Solving this problem requires a modification of the Gram-Schmidt algorithm:

1. Instead of a 12×12 ball (i.e., $n = 12$ in the original algorithm), consider an 8×8 ball by choosing to eliminate the t , p_t , p_ϕ , and S_t components. The time component ξ_t of each tangent vector is irrelevant since nothing in the problem is explicitly time dependent; the first column of the Jacobian matrix is zero, so ξ_t is not necessary to determine the time-evolution.⁷ The other three components are determined by the constraints as described above.
2. Given eight initial random tangent vectors, apply the Gram-Schmidt process to form an 8×8 ball. For each vector, determine the three missing components using the constraints, and then evolve the system using

$$\frac{dU}{d\tau} = J_\tau U$$

as before. (Now U represents a 12×8 matrix instead of a 12×12 ball.)

3. At each time T , extract the relevant eight components from each vector to form a new 8×8 ellipsoid, apply Gram-Schmidt, and then fill in the missing components using the constraints, yielding again a 12×8 matrix. The projected norms of the eight tangent vectors contribute to the running sums for the Lyapunov exponents as in the original algorithm.

The algorithm above yields eight Lyapunov exponents for the Papapetrou system of equations.

In order to implement this algorithm, we must have a method for constructing a full tangent vector ξ from an eight-component vector $\tilde{\xi}$. The method is as follows:

1. Let $\tilde{\mathbf{y}}' = \tilde{\mathbf{y}} + \epsilon \tilde{\xi}$ for a suitable choice of ϵ .
2. Fill in the missing components of $\tilde{\mathbf{y}}'$ using the constraints to form \mathbf{y}' , taking care that \mathbf{y} and \mathbf{y}' have the same constraint branches.
3. Infer the full tangent vector using $\xi = \frac{\mathbf{y}' - \mathbf{y}}{\epsilon}$.

This technique depends on the choice of ϵ , and fails when ϵ is too small or too large. Using the techniques discussed in the next section to calibrate the system, we find that $\epsilon \approx 10^{-5}$ – 10^{-6} works well in practice.

2.4.2.3 Two rigorous techniques

It should be clear from the discussion above that extracting all eight Lyapunov exponents is difficult, and in practice the techniques are finicky, depending (among other things) on the choice of ϵ as

⁷Also, the time piece is discarded in the projected norm formalism in any case (Sec. 2.4.1).

described in Sec. 2.4.2.2 above. How, then, can we be confident that the results make sense? Fortunately, there are two techniques that give rigorous Lyapunov exponents by managing to sidestep the constraint complexities entirely.

First, it is always possible to calculate the single largest exponent using the Jacobian method without considering the constraint subtleties. The complexity of the main Jacobian approach involves the competing requirements of Gram-Schmidt orthogonality and constraint satisfaction, but in the case of only one vector these difficulties vanish. Since the equations of motion preserve the constraints, an initial constraint-satisfying tangent vector retains this property throughout the integration. Thus, we begin with a vector constructed as in Sec. 2.4.2.1 and evolve it (without rescaling) along with the equations of motion. Other than the requirement of constraint satisfaction, its initial components are arbitrary, so it evolves in the direction of largest stretching and eventually points in the largest Lyapunov direction. The logarithm of its projected norm then contributes to the sum for the largest Lyapunov exponent.

Second, we can implement a deviation vector approach as described in Sec. 2.3.1. Given an initial condition \mathbf{y}_0 , we construct a nearby initial condition \mathbf{y}'_0 as in Sec. 2.4.2.1 and then evolve them both forward. In principle, an approximation for the largest Lyapunov exponent is then $\frac{1}{\tau} \log(\|\mathbf{y}' - \mathbf{y}\|/\|\mathbf{y}'_0 - \mathbf{y}_0\|) \equiv \frac{1}{\tau} \log(\|\delta\mathbf{y}\|/\|\delta\mathbf{y}_0\|)$. In practice (for chaotic systems) the method saturates: for a given initial deviation, say $\|\delta\mathbf{y}_0\| \sim 10^{-6}$, once the initial conditions have diverged by a factor of $\sim 10^6$ the method breaks down.⁸ (The traditional solution to the saturation problem is to rescale the deviation before it saturates, but such a rescaling in this case violates the constraints.) Despite its limitations, this unrescaled deviation vector technique is valuable, since it tracks the correct solution until the saturation limit is reached, and avoids the subtleties associated with the constraints.

With these two techniques in hand, we have a powerful method for verifying that the largest Lyapunov exponent produced by the Gram-Schmidt method is correct. This, in turn, gives us confidence that the other Lyapunov exponents produced by the main algorithm are meaningful as well.

2.5 Implementation details

2.5.1 Some numerical comments

Finally, we discuss some specialized issues related to integrating the Papapetrou equations on a computer. The primary subjects are the formulation of the equations, optimization techniques, and error checking.

⁸This underscores the point that chaos is essentially a *local* phenomenon. *Any* unrescaled deviation vector approach must saturate, since no bounded system can have trajectories that diverge for arbitrarily long times.

Our choice to write the Papapetrou equations using the spin vector is motivated partially by numerical considerations. The spin vector approach has nice properties compared to the tensor approach as $S \rightarrow 0$. Comparing their covariant derivatives is instructive:

$$\begin{aligned}\nabla_{\vec{v}} S_\mu &= -p_\mu \left(R^{*\alpha}{}_\beta{}^{\gamma\delta} S_\alpha v^\beta p_\gamma S_\delta \right) \\ \nabla_{\vec{v}} S^{\mu\nu} &= p^\mu v^\nu - p^\nu v^\mu = 2p^{[\mu} v^{\nu]}.\end{aligned}$$

Though simpler in form, the derivative of $S^{\mu\nu}$ has unfortunate numerical properties for small S , since in the limit $S \rightarrow 0$ we have $p^\mu \rightarrow v^\mu$: the difference $p^\mu v^\nu - p^\nu v^\mu$ goes to zero in principle but in practice is plagued by numerical roundoff errors. Since $S \ll 1$ is the most physically interesting limit, the vector approach is more convenient for our purposes.

Calculating the many tensors and derivatives which go into the Papapetrou equations and the corresponding Jacobian matrix is a considerable task. As a first step, we use GRTensor for Maple to calculate all relevant quantities, and we use Maple’s optimized C output to create C code automatically. Due to the symmetries of the Riemann tensor and the metric, many terms are identically zero, which significantly reduces the number of required operations. For example, in order to calculate $R^{*\alpha}{}_\beta{}^{\gamma\delta} S_\alpha v^\beta p_\gamma S_\delta$ we need four loops, which constitutes $4^4 = 256$ evaluations, but in fact $R^{*\alpha}{}_\beta{}^{\gamma\delta}$ has only 80 nonzero components. Performing loop unrolling by writing these terms to an optimized derivatives file consisting of explicit sums speeds up calculation by an order of magnitude compared to nested `for` loops.

Another optimization involves the choice of coordinates used in the metric, which has significant consequences for the size of the tensor files and the number of floating point operations required. Simply using $\mu = \cos \theta$ in the Kerr metric reduces the size of the Riemann derivatives by at least a factor of two.⁹ Since these derivatives are the bottleneck in the calculation of the Jacobian matrix, we can get more than a 50% improvement in performance with even this simple variable transformation.

All integrations were performed using a Bulirsch-Stoer integrator adapted from *Numerical Recipes* [28]. Occasional checks with a fifth-order Runge-Kutta integrator were in agreement. We verified the Papapetrou integration by checking errors in the constraints and conserved quantities; for an orbit such as that shown in Fig. 2.6, all errors are at the 10^{-11} level after $\tau = 10^5 M$.

As should be clear from Sec. 2.5.2 below, the Jacobian matrix of the Papapetrou equations has a large number of terms, and it is essential to verify its correctness by using a diagnostic that compares $\mathbf{Df} \cdot \delta \mathbf{y}$ with the difference $\mathbf{f}(\mathbf{y} + \delta \mathbf{y}) - \mathbf{f}(\mathbf{y})$ for a suitable constraint-satisfying $\delta \mathbf{y}$. It is not sufficient for the difference merely to be small: we must calculate the quantity $\mathbf{f}(\mathbf{y} + \delta \mathbf{y}) - \mathbf{f}(\mathbf{y}) - \mathbf{Df} \cdot \delta \mathbf{y}$ for several values of $\delta \mathbf{y}$ and verify that each component scales as $\|\delta \mathbf{y}\|^2$. An early implementation of the

⁹*Warning:* This variable substitution changes the handedness of the coordinate system, since the unit vector $\hat{\mu}$ points opposite to $\hat{\theta}$. This in turn introduces an extra minus sign in the Levi-Civita tensor $\epsilon^{\alpha\beta\gamma\delta}$, which appears many times in the Papapetrou equations and the corresponding conserved quantities. The author discovered this subtlety the hard way.

Jacobian matrix, which gave nearly identical results for $\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y})$ and $\mathbf{Df} \cdot \delta\mathbf{y}$, nevertheless had an undetected $\mathcal{O}(S^2)$ error. The unrescaled deviation vector approach showed a discrepancy with the Jacobian method,¹⁰ which showed spurious chaotic behavior. The $\|\delta\mathbf{y}\|^2$ scaling method described above eventually diagnosed the problem, which resulted from a missing term in $\partial\dot{S}_\mu/\partial S_\nu$ (Sec 2.5.2).

2.5.2 The Jacobian matrix

For reference, we write out explicit equations for part of the Jacobian matrix of the Papapetrou equations.

The Jacobian matrix of a system of differential equations, specialized to the case at hand, is as follows:

$$\begin{pmatrix} \frac{\partial \dot{x}^\mu}{\partial x^\nu} & \frac{\partial \dot{x}^\mu}{\partial p_\nu} & \frac{\partial \dot{x}^\mu}{\partial S_\nu} \\ \frac{\partial \dot{p}_\mu}{\partial x^\nu} & \frac{\partial \dot{p}_\mu}{\partial p_\nu} & \frac{\partial \dot{p}_\mu}{\partial S_\nu} \\ \frac{\partial \dot{S}_\mu}{\partial x^\nu} & \frac{\partial \dot{S}_\mu}{\partial p_\nu} & \frac{\partial \dot{S}_\mu}{\partial S_\nu} \end{pmatrix} \quad (2.45)$$

Once we calculate $\partial \dot{x}^\mu / \partial x^\nu = v^\mu_{,\nu}$, all the other derivatives can be expressed in terms of the derivatives of v^μ , the tensors and connection coefficients, and Kronecker δ s.

Written out in full, and the Papapetrou equations are as follows:

$$\dot{x}^\mu = v^\mu \quad (2.46)$$

$$\dot{p}_\mu = -R_{\mu\nu}^{*\alpha\beta} v^\nu p_\alpha S_\beta + \Gamma_{\beta\mu}^\alpha p_\alpha v^\beta \quad (2.47)$$

$$\dot{S}_\mu = -p_\mu \left(R_{\beta}^{*\alpha\gamma\delta} S_\alpha v^\beta p_\gamma S_\delta \right) + \Gamma_{\beta\mu}^\alpha S_\alpha v^\beta \quad (2.48)$$

We measure τ and r in units of M (the mass of the central body), p_μ in units of the particle rest mass μ , and S_μ in terms of the product μM . The overdot is an ordinary derivative with respect to proper time: $\dot{x} \equiv dx/d\tau$.

The unusual placement of indices on R^* is motivated by the form of the Jacobian matrix. The index placement shown above brings the equations into a form where the indices on p_μ and S_μ are always lowered, which simplifies the Jacobian matrix since (for example) $\partial p_\mu / \partial x^\mu = 0$. Otherwise the Jacobian matrix is unnecessarily complicated; for example, if p^μ appeared anywhere on the right hand side then we would have $\partial p^\mu / \partial x^\nu \neq 0$, which would contribute to J_τ .

As discussed in Sec 2.2.1, the supplementary condition $p_\mu S^{\mu\nu} = 0$ [Eq. (2.6)] leads to the equation for v^μ in terms of p^μ :

$$v^\mu = N(p^\mu + w^\mu) = N\tilde{v}^\mu, \quad (2.49)$$

¹⁰This illustrates the value of calculating the Lyapunov exponents using two different methods.

where

$$\tilde{v}^\mu = p^\mu + w^\mu \quad (2.50)$$

and

$$w^\mu = -{}^*R^{*\mu\alpha\beta\gamma}S_\alpha p_\beta S_\gamma. \quad (2.51)$$

N is a normalization constant fixed by $v_\mu v^\mu = -1$.

The calculation of the partial derivatives \dot{x}^μ in Eq. (2.45) proceeds as follows. From the relation for $v^\mu = N\tilde{v}^\mu$, we have

$$\frac{\partial \dot{x}^\mu}{\partial x^\nu} = v^\mu_{,\nu} = N\tilde{v}^\mu_{,\nu} + N_{,\nu}\tilde{v}^\mu.$$

Now, $\tilde{v}^\mu_{,\nu} = p^\mu_{,\nu} + w^\mu_{,\nu} = p_\alpha g^{\alpha\mu}_{,\nu} - {}^*R^{*\mu\alpha\beta\gamma}_{,\nu}S_\alpha p_\beta S_\gamma$, so the first term is easy. The second term is trickier: from the expression for v^μ , we have that $-1 = v^\mu v_\mu = N^2(p^\mu p_\mu + 2w^\mu p_\mu + w^\mu w_\mu) = N^2(-1 + 2w^\mu p_\mu + w^\mu w_\mu)$, so we have

$$N = (1 - 2w^\mu p_\mu - w^\mu w_\mu)^{-1/2}.$$

Differentiating gives

$$\begin{aligned} N_{,\nu} &= N^3 (p_\alpha w^\alpha_{,\nu} + w^\alpha_{,\nu} w_\alpha + \tfrac{1}{2} w^\alpha w^\beta g_{\alpha\beta,\nu}) \\ &= N^3 (\tilde{v}_\alpha w^\alpha_{,\nu} + \tfrac{1}{2} w^\alpha w^\beta g_{\alpha\beta,\nu}), \end{aligned}$$

where we have re-labeled the dummy index ($\mu \rightarrow \alpha$). Summing the various terms, we have

$$\begin{aligned} v^\mu_{,\nu} &= N [p_\alpha g^{\alpha\mu}_{,\nu} + w^\mu_{,\nu} \\ &\quad + v^\mu (v_\alpha w^\alpha_{,\nu} + \tfrac{1}{2} N w^\alpha w^\beta g_{\alpha\beta,\nu})]. \end{aligned} \quad (2.52)$$

The expression for $\partial \dot{x}^\mu / \partial p_\nu$ is similar to $v^\mu_{,\nu}$, but it is simpler because the derivative of the metric with respect to the momentum is zero. As before, we use the product rule:

$$\frac{\partial v^\mu}{\partial p_\nu} = N \frac{\partial \tilde{v}^\mu}{\partial p_\nu} + \frac{\partial N}{\partial p_\nu} \tilde{v}^\mu.$$

The first term requires

$$\frac{\partial \tilde{v}^\mu}{\partial p_\nu} = \frac{\partial p^\mu}{\partial p_\nu} + \frac{\partial w^\mu}{\partial p_\nu} = g^{\mu\nu} - {}^*R^{*\mu\alpha\nu\beta}S_\alpha S_\beta \equiv g^{\mu\nu} + W^{\mu\nu}.$$

Note that $W^{\mu\nu}$ is symmetric. The second term requires

$$\frac{\partial N}{\partial p_\nu} = N^3 (W^{\alpha\nu} p_\alpha + w^\alpha \delta_\alpha^\nu + W^{\alpha\nu} w_\alpha) = N^3 (w^\nu + \tilde{v}_\alpha W^{\alpha\nu}).$$

Summing the terms gives

$$\frac{\partial \dot{x}^\mu}{\partial p_\nu} = N(g^{\mu\nu} + W^{\mu\nu} + Nv^\mu w^\nu) + Nv^\mu v_\alpha W^{\alpha\nu} \quad (2.53)$$

with

$$\frac{\partial w^\mu}{\partial p_\nu} \equiv W^{\mu\nu} = -{}^*R^{*\mu\alpha\nu\beta} S_\alpha S_\beta. \quad (2.54)$$

Finally, we calculate $\partial \dot{x}^\mu / \partial S_\nu$. With

$$\frac{\partial \tilde{v}^\mu}{\partial S_\nu} = -S_\alpha p_\beta ({}^*R^{*\mu\alpha\beta\nu} - {}^*R^{*\mu\nu\alpha\beta}) \equiv V^{\mu\nu},$$

and

$$\frac{\partial N}{\partial S_\nu} \tilde{v}^\mu = Nv^\mu v_\alpha V^{\alpha\nu},$$

we have

$$\frac{\partial \dot{x}^\mu}{\partial S_\nu} = NV^{\mu\nu} + Nv^\mu v_\alpha V^{\alpha\nu}. \quad (2.55)$$

We calculate the derivatives of \dot{p}_μ and \dot{S}_μ using $v^\mu_{,\nu}$, the product rule, and the derivatives of the various tensors in the problem. The full results appear in the appendix.

2.6 Integrability and chaos

2.6.1 Phase space and constants of the motion

Having laid the foundation for the numerical calculation of Lyapunov exponents, we now discuss some general aspects of dynamical systems relevant to our study. A dynamical system with n coordinates has a $2n$ -dimensional phase space, typically consisting of generalized positions and their corresponding conjugate momenta. Motion in the phase space is arbitrary in general, but when there are integrals of the motion then the flow is confined to a surface on which the integral is constant. This can be seen most easily by transforming to angle-action coordinates, where the surface is an invariant (multidimensional) torus.

A system with n coordinates and n constants of the motion is *integrable* and cannot have chaos (though the motion can still be quasiperiodic or exhibit other complicated behavior). For example, we can consider geodesic orbits around a Kerr black hole to have eight degrees of freedom ($n = 4$) and four constants of the motion—particle rest mass μ , energy E , axial or z angular momentum L_z , and Carter constant Q —which are enough to integrate the equations of motion explicitly. Alternatively, we may look at Kerr spacetime as having a 6-dimensional phase space by eliminating time (which is simply a reparameterization of the proper time) and using rest mass conservation to eliminate one momentum coordinate. Then the three integrals E , L_z , and Q are sufficient to integrate the motion.

(In practice, we allow all four momenta to evolve freely; the normalization is then a constraint which can be checked for consistency at the end of the integration.)

In the case of a spinning test particle, the extra spin degrees of freedom create the possibility for chaotic behavior. Moreover, since Q is not conserved in the case of nonzero spin, even without the extra spin degrees of freedom the potential for chaos would exist. Kerr spacetime has just enough constants to make the system integrable; losing Q reduces the number of analytic integrals below the critical level required to guarantee integrability.¹¹

2.6.2 Hamiltonian systems

2.6.2.1 Lyapunov exponents for Hamiltonian flows

The phase space flow of Hamiltonian systems is constrained by more than the integrals of the motion. In particular, the Lyapunov exponents of a Hamiltonian system come in pairs $\pm\lambda$; i.e., if λ is a Lyapunov exponent then so is $-\lambda$ [26]. Geometrically, this means that if one semimajor axis of the phase-space ellipsoid stretches an amount $e^{\lambda\tau} = L$, another axis must shrink by an amount $e^{-\lambda\tau} = 1/L$. One consequence of this property is that the product of the lengths of the axes is 1. Since the ellipsoid volume is proportional to this invariant product, Liouville's theorem on the conservation of phase space volume follows as a corollary.

The $\pm\lambda$ property of Hamiltonian flows results from the symplectic nature of the Jacobian matrix for Hamiltonian dynamical systems.¹² But a naïve analysis of the Jacobian matrix of the Papapetrou equations shows that it is not symplectic in the canonical sense. Nevertheless, the Papapetrou equations can be derived from a Lagrangian [30], and can be cast in Hamiltonian form by use of a free Hamiltonian with added constraints (following the method of Dirac [31] as discussed in [32]). As a consequence, we could in principle find coordinates in which the Jacobian matrix is symplectic with respect to the canonical symplectic matrix. Fortunately, this is an unnecessary complication, since the underlying dynamics are independent of the coordinates.

2.6.2.2 Exponents for spinning test particles

As discussed in Sec. 2.4.2.2, the lack of explicit time dependence and the three constraints reduce the degrees of freedom from twelve to eight, which leaves the possibility of eight nonzero Lyapunov exponents. The phase space flow is further constrained by the constants of the motion, energy and z angular momentum; corresponding to each constant should be a zero Lyapunov exponent, since trajectories that start on an invariant torus must remain there. This leaves

¹¹It is possible that deformations of Kerr geometry that destroy Q nevertheless possess a numerical integral that preserves integrability, in analogy with some galactic potentials [29], but the loss of Q certainly ends the *guarantee* of integrability.

¹²A matrix S is symplectic with respect to the canonical symplectic matrix J if $S^T J S = J$, where $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ and I is the $n \times n$ identity matrix.

six exponents potentially nonzero. Since the exponents must come in pairs $\pm\lambda$, there should be at most three independent nonzero exponents.

2.7 Results

First we give results for the dynamics of the Papapetrou equations in the extreme (and unphysical) limit $S = 1$, which represents a violation of the test-particle approximation but is still mathematically well-defined. We find the presence of chaotic orbits (in agreement with [3]). We next examine the effects of varying S , including the limit $S \ll 1$. Finally, we investigate more thoroughly the dynamics for physically realistic spins.

2.7.1 Chaos for $S = 1$

2.7.1.1 Maximally spinning Kerr spacetime

In a background spacetime of a maximally spinning Kerr black hole ($a = 1$) there are unambiguous positive Lyapunov exponents for a range of physical parameters when $S = 1$. We show a typical orbit that produces nonzero Lyapunov exponents in Fig. 2.6. The orbit has energy $E = 0.8837\mu$, z angular momentum $J_z = 2.0667\mu M$, and the radius ranges from pericenter $r_p = 1.7M$ to apocenter $r_a = 6.7M$. The Lyapunov integrations typically run for $10^4 M$, which corresponds approximately to 400 ϕ -orbital periods.

We can illustrate the presence of a chaotic orbit by plotting the natural logarithm of the i th ellipsoid axis $\log[r_i(\tau)]$ vs. τ [Eq. (2.40)], so that the slope is the Lyapunov exponent, as shown in Fig. 2.7.¹³ There appear to be two nonzero Lyapunov exponents; the third largest exponent is consistent with zero, as shown in Fig. 2.8. The reflection symmetry of the figure is a consequence of the exponent pairing: for each line with slope λ , there is a second line with slope $-\lambda$.

The main plot in Fig. 2.7(a) is generated by the modified Gram-Schmidt algorithm (Sec. 2.4.2.2). Recall that this method depends on the value of ϵ used to infer the tangent vector; we find a valid ϵ by calibrating it using the rigorous Jacobian method, which must yield an exponent that matches the largest exponent from the modified Gram-Schmidt method. The plot in Fig. 2.7(a) represents the case $\epsilon = 10^{-6}$; it is apparent that the two methods agree closely. The unrescaled deviation vector method provides an additional check on the validity of the largest exponent, as shown in Fig. 2.7(b). As expected, the unrescaled approach closely tracks the full Jacobian approach until it saturates.

The numerical values of the exponents are shown in Table 2.1. The $\pm\lambda$ property is best satisfied by $\pm\lambda_{\max}$, the exponents with the largest absolute value. The exponents are least-squares fits to

¹³It is traditional to plot $\log[r_i(\tau)]/\tau$, which converges to the Lyapunov exponent as $\tau \rightarrow \infty$, but it is much easier to identify the linear growth of $\log[r_i(\tau)]$ than to identify the convergence of $\log[r_i(\tau)]/\tau$. The $\pm\lambda$ property is also clearer on such plots.

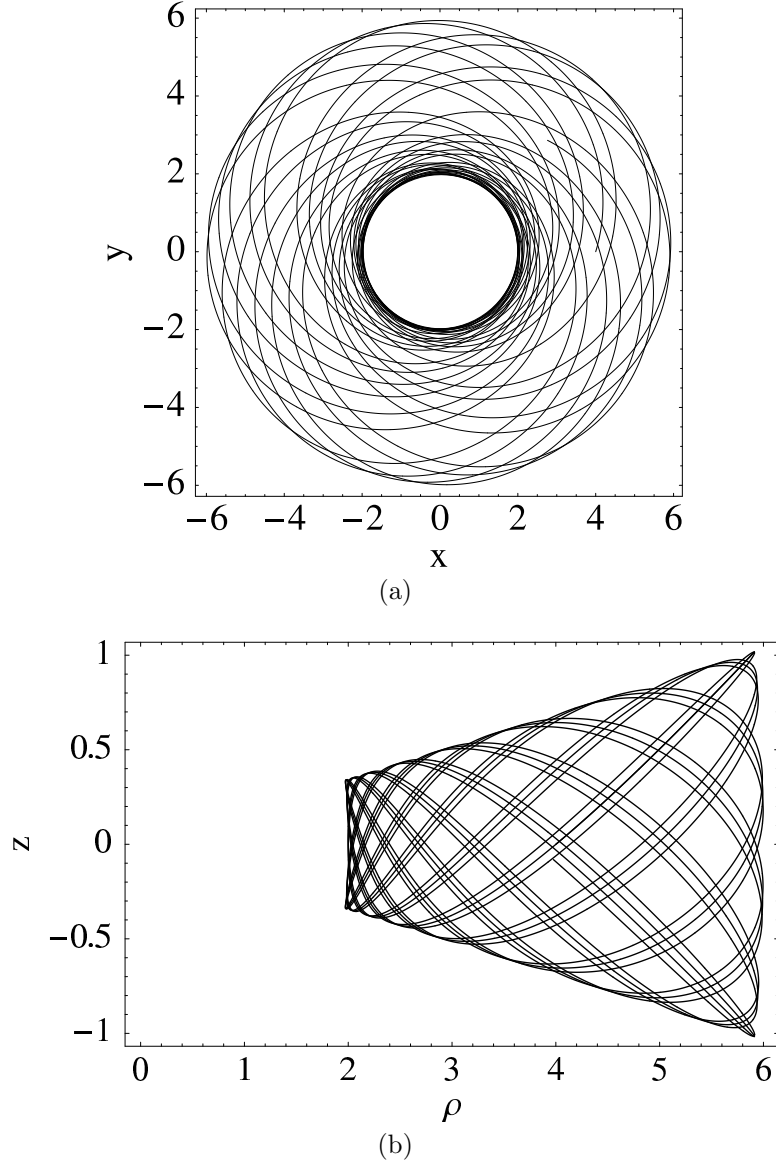


Figure 2.5: The orbit of a non-spinning ($S = 0$) test particle in maximal ($a = 1$) Kerr spacetime, plotted in Boyer-Lindquist coordinates. (a) $y = r \sin \theta \sin \phi$ vs. $x = r \sin \theta \cos \phi$; (b) z vs. $\rho = \sqrt{x^2 + y^2}$. The orbital parameters are $E = 0.8837 \mu$ and $J_z = 2.0667 \mu M$, with pericenter $2.0 M$ and apocenter $6.0 M$.

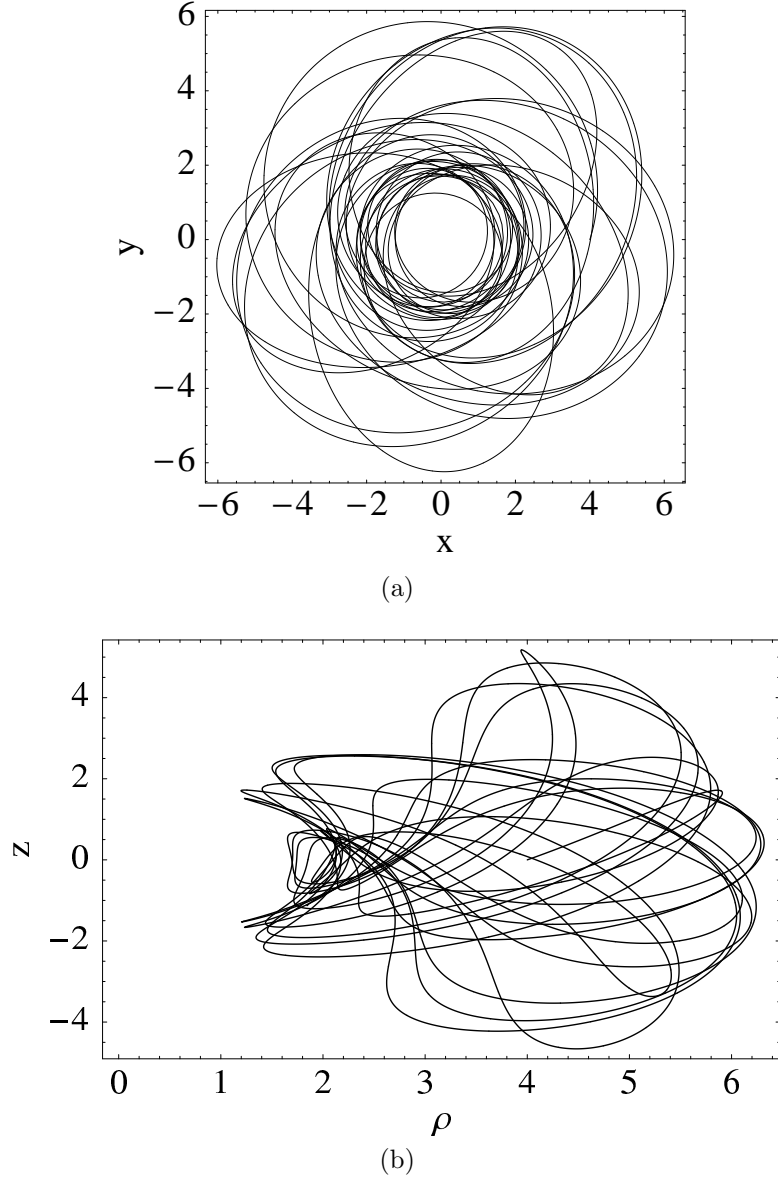


Figure 2.6: The orbit of a maximally spinning ($S = 1$) particle in maximal Kerr spacetime, for $E = 0.8837 \mu$ and $J_z = 2.0667 \mu M$ (the same values as in Fig. 2.5 above). The spin has initial values of $S^{\hat{r}} = S^{\hat{\theta}} = 0.1$, corresponding to an initial angle of 54° with respect to the vertical in a fiducial (ZAMO) rest frame. As in Fig. 2.5, we plot y vs. x in (a) and z vs. ρ in (b). The spin causes significant deviations from geodesic orbits.

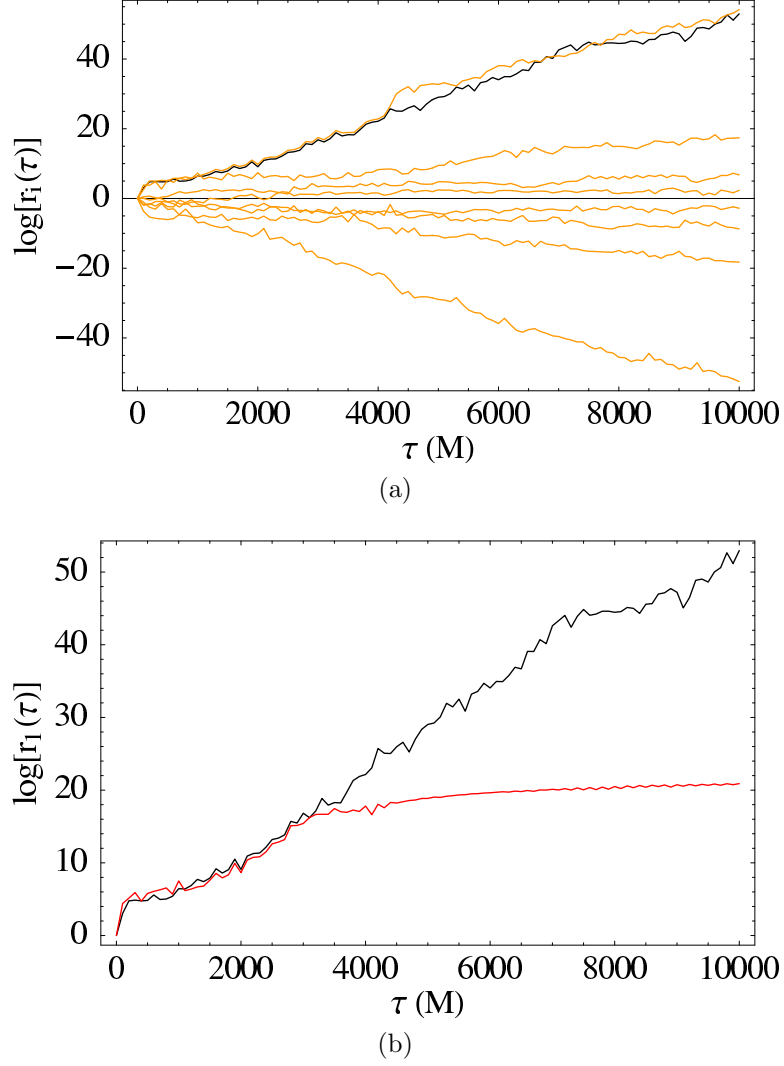


Figure 2.7: Natural logarithms of the phase space ellipsoid axes vs. proper time in Kerr spacetime with $S = 1$. The slopes of the lines are the Lyapunov exponents; the largest exponent is approximately $\lambda_{\max} = 5 \times 10^{-3} M^{-1}$. The initial conditions are the same as in Fig. 2.6, and one point is recorded at each time $T = 100 M$ (Sec. 2.3.2.1). (a) Full Gram-Schmidt Jacobian method (light) with rigorous Jacobian method (dark). The full GS method is rescaled at each time T according the algorithm in Sec. 2.3.2, while the rigorous Jacobian method is unrescaled. The two methods agree closely on the value of the largest Lyapunov exponent. (b) Rigorous Jacobian method compared to unrescaled deviation vector method. Note that the latter method, which started with a deviation of size 10^{-7} , saturates at ~ 16 . This corresponds to a growth of $e^{16} \approx 9 \times 10^6$, which means that the separation has grown to a size of order unity.

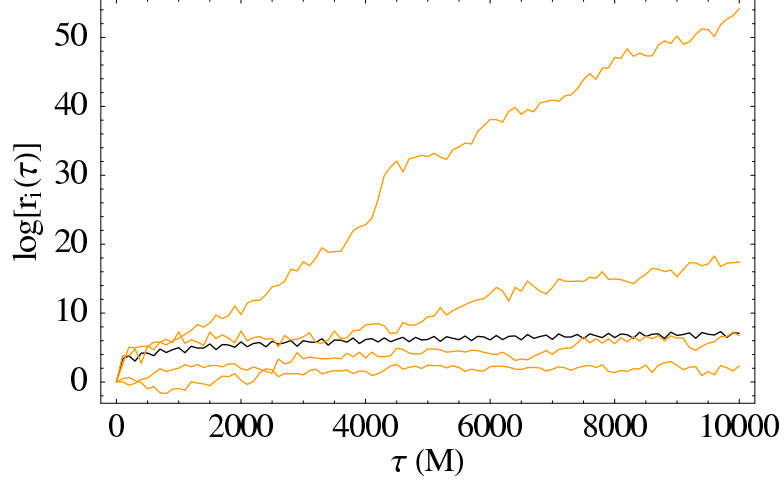


Figure 2.8: Ellipsoid axis lengths from the upper half of Fig. 2.7(a) (light), compared to an integration with zero spin and hence zero Lyapunov exponent (dark). Only two of the four lines represent exponents distinguishable from zero.

the data, with approximate standard errors of 1%. These errors are not particularly meaningful since the exponents themselves can vary by $\sim 10\%$ depending on the initial direction of the deviation vector. Moreover, even exponents that appear nonzero may be indistinguishable from zero in the sense of Fig. 2.8; for such exponents a “1%” error on the fit is meaningless.

For initial conditions considered in Fig. 2.6, and other orbits in the strongly relativistic region near the horizon, typical largest Lyapunov exponents are on the order of a few $\times 10^{-3}/M$. For the particular case illustrated in Fig. 2.6, we have $\lambda_{\max} \approx 5 \times 10^{-3} M^{-1}$, which implies an e -folding timescale of $\tau_{\lambda} \equiv 1/\lambda \approx 2 \times 10^2 M$. This is strongly chaotic, with a significant divergence in approximately eight ϕ -orbital periods.

Based on integrations in the case of zero spin, which corresponds to no chaos (Lyapunov exponents all zero), we can determine how quickly the exponents approach zero numerically.¹⁴ Fig. 2.8 compares the four apparently positive exponents with a known zero exponent. Only two of the four exponents are unambiguously distinguishable from zero, consistent with the argument in Sec. 2.6.2 that there should be at most three independent nonzero exponents.

Finally, we note that the components of the direction of largest stretching are all nonzero in general. The chaos is not confined to the spin variables alone, but rather mixes all directions. This indicates that chaos could in principle manifest itself in the gravitational waves from extreme mass-ratio binaries—but see Sec. 2.7.3 below.

¹⁴As noted in the introduction, it is possible for integrable but unstable orbits to have positive Lyapunov exponents. We avoid this issue by choosing a baseline orbit that is not unstable.

Table 2.1: Lyapunov exponents in Kerr spacetime in units of $10^{-3} M^{-1}$, using a least squares fit. The exponents correspond to the semimajor axis evolution shown in Fig. 2.7(a). As is typical with the Gram-Schmidt Jacobian method, the exponents with the largest magnitudes are determined most accurately, and thus show the $\pm\lambda$ property most clearly. The standard errors on the fit are $\sim 1\%$ for each exponent, but these errors are dominated by two systematic errors: (1) the variation due to different choices of initial (random) tangent vectors; (2) nonzero numerical values even for exponents that converge to zero eventually. In particular, the four smallest exponents (in absolute value) are indistinguishable from zero (see Fig. 2.8).

$+\lambda$	5.5	1.5	0.56	0.25
$-\lambda$	5.3	1.6	0.76	0.072

Table 2.2: Lyapunov exponents in Schwarzschild spacetime in units of $10^{-3} M^{-1}$, using a least squares fit. The exponents correspond to the semimajor axis evolution shown in Fig. 2.10(a), which is similar to the orbit in Fig. 4(d) of Ref. [3]. As with the Kerr case (Table 2.1), the standard errors on the fit are $\sim 1\%$ for each exponent, and the same caveats apply. The four smallest exponents (in absolute value) are indistinguishable from zero in the sense of Fig. 2.8.

$+\lambda$	1.2	0.67	0.21	0.0063
$-\lambda$	1.5	0.57	0.10	0.00023

2.7.1.2 Schwarzschild spacetime revisited

We now reconsider the case of a spin $S = 1$ particle in Schwarzschild spacetime, as investigated in Ref. [3]. Fig. 2.9 shows an orbit similar to a chaotic orbit considered there (Fig. 4(d) in [3]). A plot of $\log[r_i(\tau)]$ vs. τ (Fig. 2.10) shows behavior similar to that in Fig. 2.7. In particular, the $\pm\lambda$ symmetry is present, apparently with two positive exponents. (The other lines are indistinguishable from zero, again using $S = 0$ orbits as a baseline.) The largest exponent of $1.5 \times 10^{-3} M^{-1}$ agrees closely with the value from Ref. [3], which reported an exponent of $\sim 2 \times 10^{-3} M^{-1}$ for a similar orbit. (This agreement is somewhat surprising, since [3] appears not to have taken the constrained nature of the deviation vectors into account. Luckily, the exponents are robust, and even unconstrained deviation vectors give nearly correct results.)

2.7.1.3 Kerr and Schwarzschild compared

The Kerr and Schwarzschild Lyapunov exponents of the previous two sections are not all that different; both are 10^{-2} – $10^{-3} M^{-1}$ in order of magnitude. Nevertheless, the two systems prove to be quite different: chaotic orbits are easy to find in Kerr spacetime for nearly any initial condition that explores the strongly relativistic region near the horizon, whereas nearly all analogous orbits in Schwarzschild spacetime are not chaotic.

Fig. 2.11 compares Kerr and Schwarzschild orbits with the same inclination angle $\iota = 10^\circ$ and eccentricity $e = 0.5$ but varying pericenters r_p . (Details of this parameterization method, mentioned

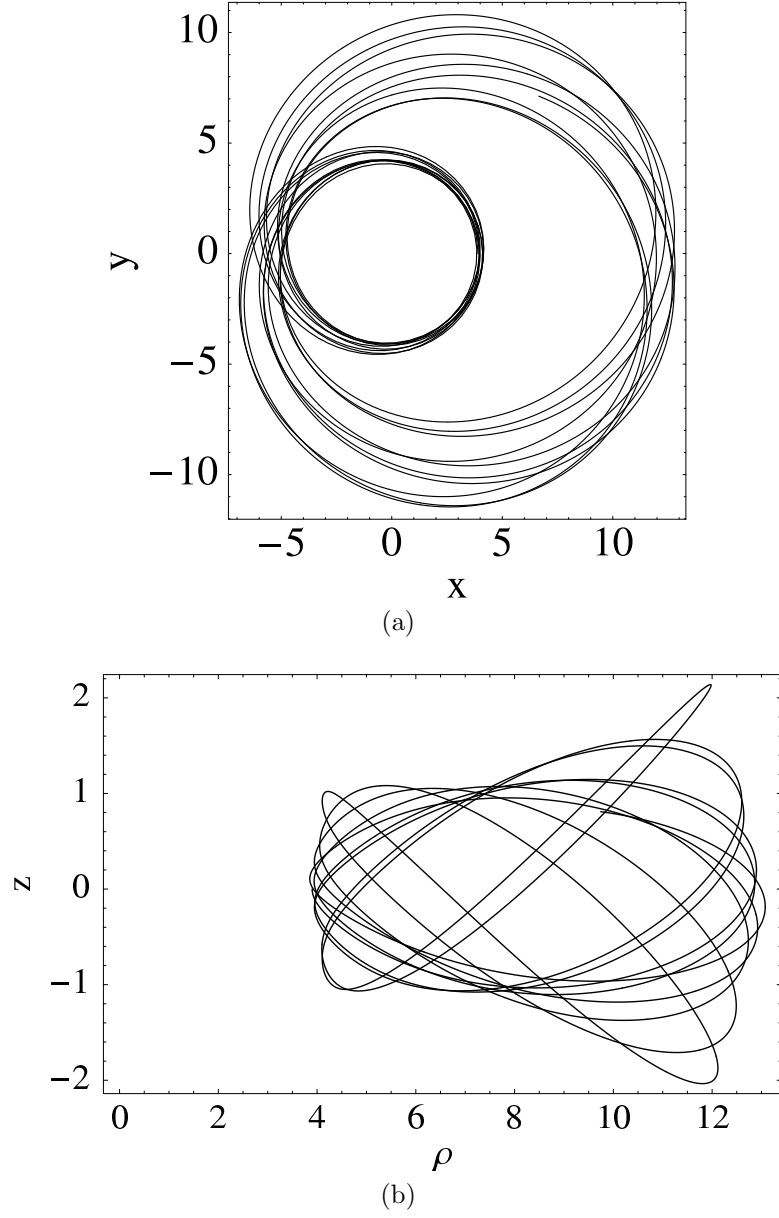


Figure 2.9: The orbit of a maximally spinning ($S = 1$) test particle in Schwarzschild spacetime for $E = 0.94738162\mu$ and $J_z = 4.0\mu M$. As before, we plot (a) y vs. x and (b) z vs. $\rho = \sqrt{x^2 + y^2}$.

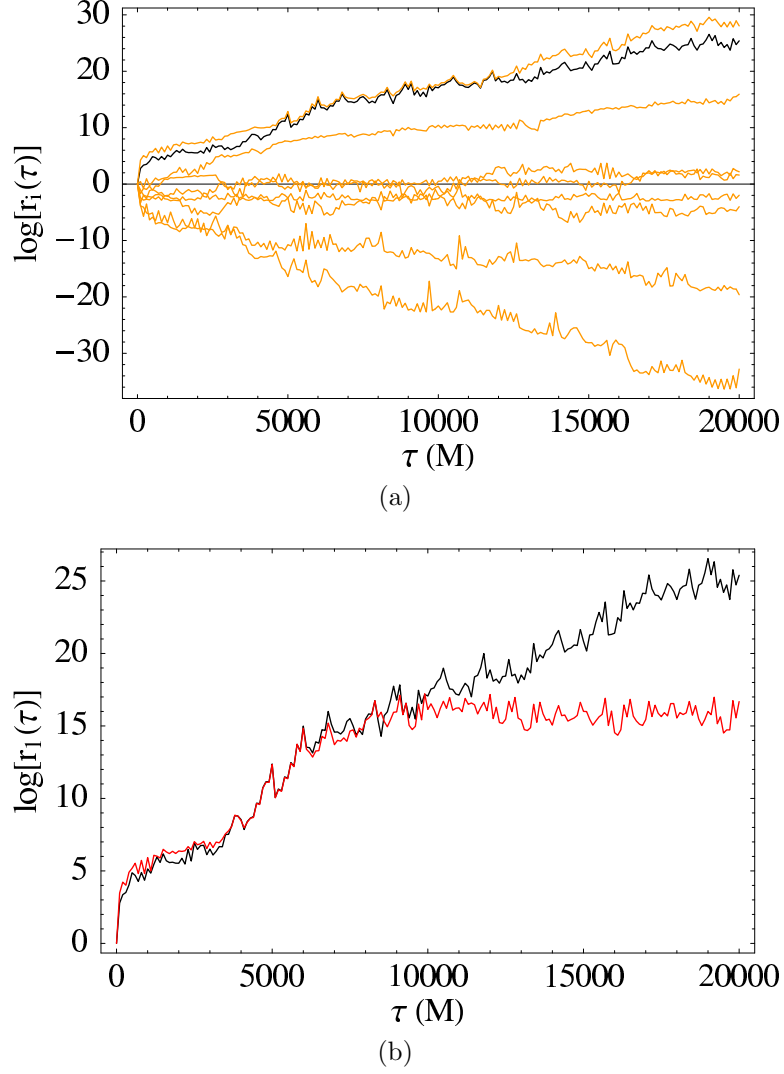


Figure 2.10: Natural logarithms of the phase space ellipsoid axes vs. proper time in Schwarzschild spacetime with $S = 1$. The largest exponent is $\lambda_{\max} \approx 1.2 \times 10^{-3} M^{-1}$. The initial conditions are the same as in Fig. 2.9. (a) Full Gram-Schmidt Jacobian method (light) with rigorous Jacobian method (dark). (b) Rigorous Jacobian method compared to unrescaled deviation vector method. As in Fig. 2.7(b), the unrescaled method eventually saturates.

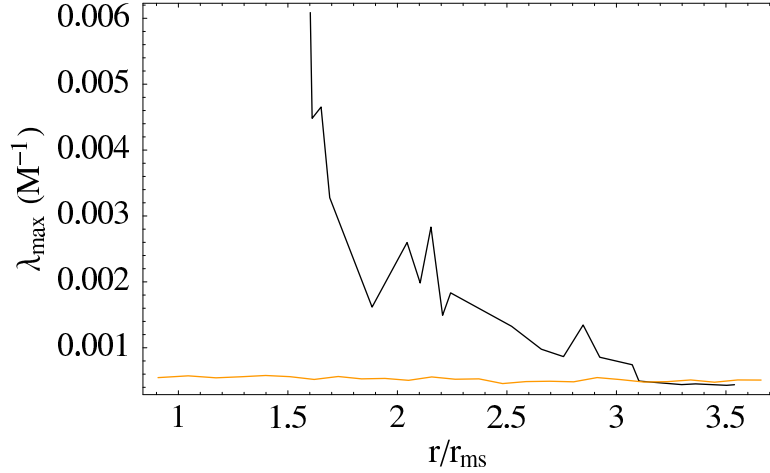


Figure 2.11: Comparison of maximally spinning ($S = 1$) particle orbits in Kerr (dark) and Schwarzschild (light) spacetime. We plot the largest Lyapunov exponent versus pericenter (normalized by the marginally stable radius). The Kerr initial conditions for the innermost orbits are essentially as in Fig. 2.6. The Schwarzschild orbits are identical to their Kerr counterparts in inclination (10°) and eccentricity ($e = 0.5$) but have the Kerr parameter a set to zero. The Schwarzschild orbits have exponents indistinguishable from zero over the entire range of parameters.

above in Sec. 2.2.3, appear in Chapter 3.) We insure that the systems are analogous by using orbits of $S = 1$ particles with the same values of r_p/r_{ms} , where r_{ms} is the radius of the marginally stable orbit in the corresponding $S = 0$ (geodesic) case. We use a Kerr geodesic integrator developed by Scott Hughes [33] to find r_{ms} , which is the smallest pericenter that still yields a stable orbit. For the values of ι and e considered, $r_{\text{ms}} = 1.0 M$ for Kerr and $r_{\text{ms}} = 4.67 M$ for Schwarzschild.

It is evident from Fig. 2.11 that the Kerr orbits are chaotic for a broad range of pericenters, with the maximum Lyapunov λ_{max} generally decreasing as the pericenter increases. In contrast, the Schwarzschild orbits are not chaotic anywhere over the entire range of valid initial conditions. In fact, we are unable to find any chaotic orbits in Schwarzschild spacetime other than the types identified by Suzuki and Maeda [3], which were exceptional cases of orbits on the edge of a generalized effective potential. In Kerr, on the other hand, chaotic orbits appear to be the rule for pericenters near r_{ms} .

2.7.2 Dependence on S

Since chaos must disappear as $S \rightarrow 0$, we expect to see the largest Lyapunov exponent approach zero in this limit. This is indeed the case: in Fig. 2.12, which shows the variation of λ_{max} with S for two different orbits, we see that the chaos unambiguously present when $S = 1$ is not present for smaller values of S . In particular, the largest Lyapunov exponent is indistinguishable from zero over the entire range $10^{-6} \leq S \leq 10^{-1}$. (The far left of the plots have data points for each decade in this range.)

Although the strength of the chaos generally decreases with S , one remarkable feature of Fig. 2.12(a)

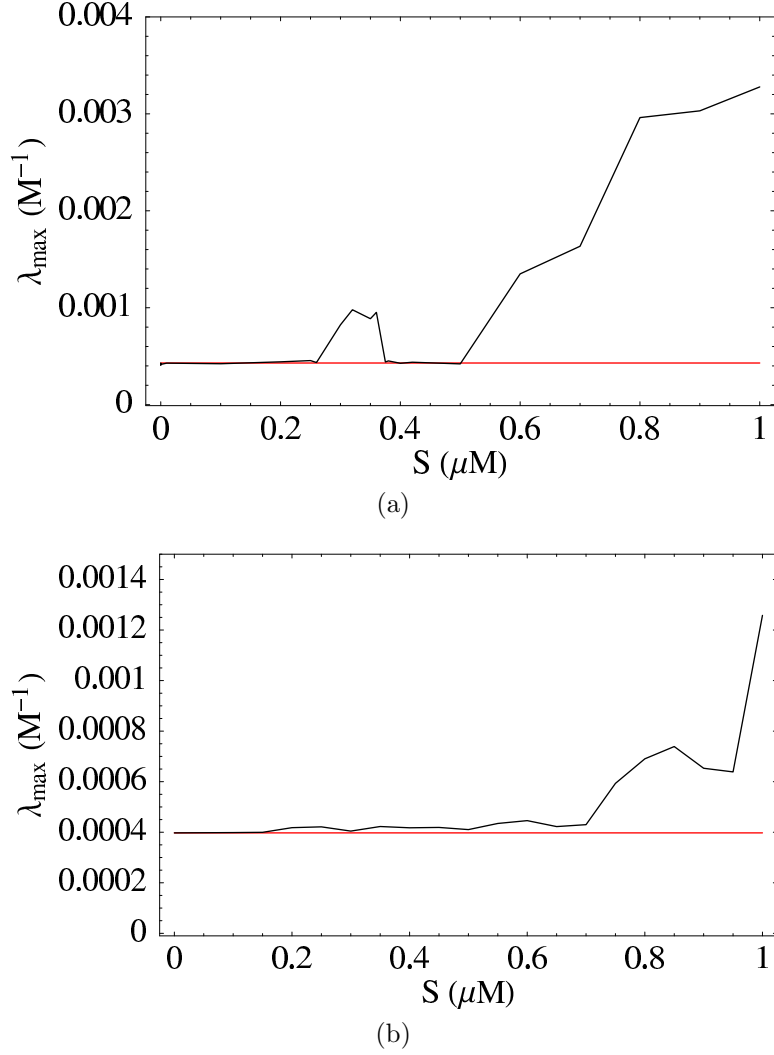


Figure 2.12: Variation of largest Lyapunov exponent vs. S . (a) The spin $S = 1$ initial conditions are the same as in Fig. 2.6. (b) Another $S = 1$ case with a different inclination angle (20°) and pericenter ($2.5 M$). As the spin decreases, we hold fixed the Kerr orbital parameters: inclination angle, eccentricity, and pericenter. Note that in (a) the chaos disappears below $S \sim 0.5$, but returns in a region centered on $S \sim 0.3$. The horizontal line in both plots is the value of λ_{\max} calculated for the baseline $S = 0$ orbit. In both (a) and (b) the Lyapunov exponent is indistinguishable from zero for physically realistic spins.

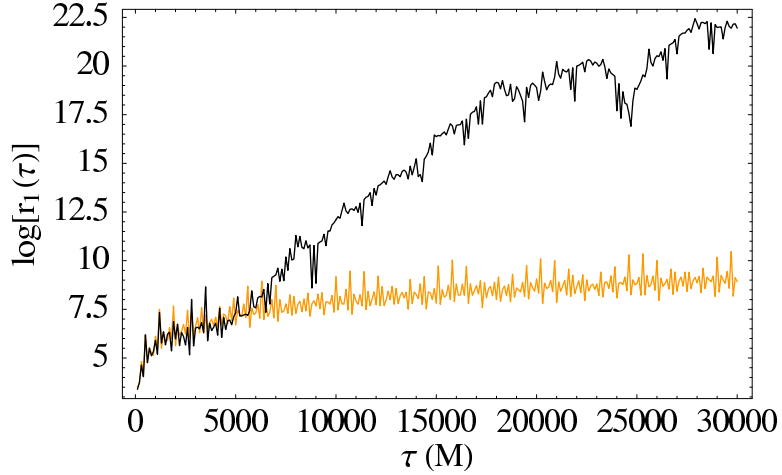


Figure 2.13: Two orbits from the “bump” in Fig. 2.12(a). The $S = 0.4$ orbit (light) is not chaotic, but the $S = 0.3$ orbit (dark) is chaotic, despite having a smaller value of the spin.

is the return of chaotic orbits between $S \sim 0.25$ and 0.4 after their disappearance at $S \sim 0.5$. The effect is qualitatively clear in Fig. 2.13. This chaotic “bump” in λ_{\max} vs. S illustrates an important theme in nonlinear dynamical systems: the *only* way to determine whether an orbit is chaotic is to do the calculation. Though we certainly expect the strength of chaos to be smaller for $S \ll 1$ than for $S \approx 1$, it is impossible, in general, to determine *a priori* whether a particular set of parameters will lead to chaotic behavior.

2.7.3 Physically realistic spins

The Papapetrou equations are only realistic in the test-particle limit, so physically realistic spins must satisfy $S \ll 1$ (Sec. 2.2.2). This corresponds to likely sources of gravitational waves for LISA [34–36], e.g., maximally spinning $\mu = 10 M_{\odot}$ black holes spiraling into supermassive $M = 10^6 M_{\odot}$ Kerr black holes, which have spin parameters of $S = \mu/M = 10^{-5}$. Because of their likely importance as emitters of gravitational waves, it is essential to understand the dynamics of such systems.

2.7.3.1 Vanishing Lyapunov exponents

We would like to be able to make a definitive statement about the presence or absence of chaos for “small” spins, e.g., values of S in the range 10^{-2} – 10^{-6} . Unfortunately, when determining Lyapunov exponents numerically, it is impossible to conclude definitively that an orbit is or is not chaotic, since to do so would require an infinite-time integration. On the other hand, for suspected non-chaotic orbits, we can provide an approximate bound on the e -folding timescale.

The numerical values of exponents suspected to be zero depend strongly on the time of the

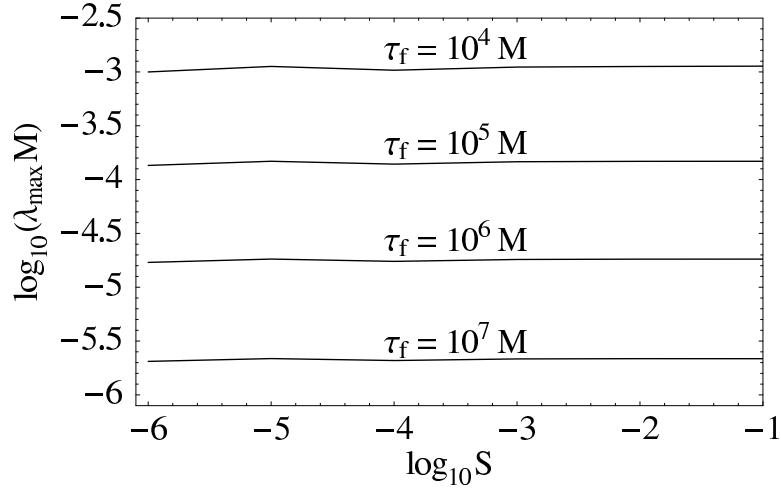


Figure 2.14: The variation of the dimensionless quantity $\lambda_{\max} M$ with final integration time τ_f for spin parameter S in the range $10^{-2} \leq S \leq 10^{-6}$. From top to bottom, the total integration time varies from $10^4 M$ to $10^7 M$. It is likely that the true Lyapunov exponent is zero.

integration. For example, for values of S in the range $10^{-2} \leq S \leq 10^{-6}$, the exponent in Fig. 2.12 appears to be $\lambda_{\max} \approx 5 \times 10^{-4} M^{-1}$, but this plot represents an integration time of only $10^4 M$. Longer integration times give correspondingly smaller estimates for the suspected zero exponents (Fig. 2.14). For the system shown in Fig. 2.12, an integration of $10^7 M$ yields an estimate of $\lambda_{\max} \approx 3.0 \times 10^{-7} M^{-1}$ for all spins in the range $10^{-2} \leq S \leq 10^{-7}$. In this case, the relevant Lyapunov timescales are at least $3 \times 10^6 M$, and are probably much longer; the size of the bound is limited only by our patience and computer budget. It seems highly likely that such orbits are not chaotic.

2.7.3.2 Spin-induced phase differences

Even if their Lyapunov exponents are zero, small spins affect the relative phase of the orbits, and since phase differences accumulate secularly [37], the spin can still affect the gravitational wave signal. It is therefore useful to have a sense of the orders of magnitude of such spin-induced phase-shifts. Tables 2.3 and 2.4 show typical values for the phase difference $\Delta\phi = \phi_{\text{geodesic}} - \phi_{\text{spin}}$ for $S = 10^{-5}$, where the geodesic and spin systems start with the same initial 4-velocity v^μ . The most useful quantity in practice is the phase shift as measured by observers at infinity, so we integrate in terms of the Boyer-Lindquist coordinate time t in place of τ . (This involves multiplying the differential equations by $d\tau/dt$ at each time step.) As is apparent from the tables, the phase shifts range broadly, from 10^{-1} to 10^{-5} radians after 2000 radial orbital periods, but tend to decrease in magnitude with increasing inclination angle or pericenter.

Reference [36] shows that the number of orbital periods in a full inspiral from $r \approx 4 M$ to the

Table 2.3: Phase shifts $\Delta\phi = \phi_{\text{geodesic}} - \phi_{\text{spin}}$ in radians as a function of orbital inclination angle ι and pericenter r_p for $a = 0.5$ and $S = 10^{-5}$. Inclination angle $\iota = 0^\circ$ is prograde equatorial and $\iota = 180^\circ$ is retrograde equatorial. The geodesic orbits and their corresponding spin orbits start with the same initial 4-velocity v^μ , and the integrations are performed using Boyer-Lindquist coordinate time t , with $t_{\text{max}} \approx (2000 \text{ times the average radial orbital period})$. The pericenters are scaled by the marginally stable radius r_{ms} , and we start at $r_p/r_{\text{ms}} = 1.5$ to guarantee the existence of valid initial conditions for the non-geodesic orbit. The spin has fixed initial values of $S^{\hat{r}} = S^{\hat{\mu}} = 0.1 S$ (with hats indicating an orthonormal basis), corresponding to initial angles of 9° to 30° with respect to the vertical in a fiducial (ZAMO) rest frame, increasing with decreasing pericenter.

ι	r_p/r_{ms}			
	1.5	2.0	2.5	3.0
10°	1.50×10^{-2}	5.69×10^{-3}	4.32×10^{-3}	2.13×10^{-3}
45°	2.79×10^{-2}	1.23×10^{-2}	1.01×10^{-2}	4.34×10^{-3}
85°	4.36×10^{-2}	2.92×10^{-3}	1.48×10^{-3}	8.24×10^{-4}
135°	-9.02×10^{-3}	-6.25×10^{-3}	-2.34×10^{-3}	-1.30×10^{-3}
170°	8.40×10^{-4}	2.85×10^{-4}	1.84×10^{-4}	7.31×10^{-5}
ι	3.5	4.0	4.5	5
10°	2.02×10^{-3}	1.27×10^{-3}	1.14×10^{-3}	7.77×10^{-4}
45°	4.60×10^{-3}	2.24×10^{-3}	1.66×10^{-3}	1.83×10^{-3}
85°	1.00×10^{-3}	2.20×10^{-3}	1.86×10^{-3}	1.26×10^{-3}
135°	-9.02×10^{-3}	-6.25×10^{-3}	-2.34×10^{-3}	-1.30×10^{-3}
170°	8.40×10^{-4}	2.85×10^{-4}	1.84×10^{-4}	7.31×10^{-5}

Table 2.4: Phase shifts $\Delta\phi = \phi_{\text{geodesic}} - \phi_{\text{spin}}$ in radians as a function of orbital inclination angle ι and pericenter r_p for $a = 1$ and $S = 10^{-5}$. As in Table 2.3, the pericenters are scaled by r_{ms} , and the spin has fixed initial values of $S^{\hat{r}} = S^{\hat{\mu}} = 0.1 S$ (corresponding in this case to initial angles of 28° to 61° , again decreasing with increasing pericenter).

ι	r_p/r_{ms}			
	1.5	2.0	2.5	3.0
10°	7.21×10^{-2}	4.58×10^{-2}	2.41×10^{-2}	1.83×10^{-2}
45°	2.37×10^{-1}	5.56×10^{-2}	2.59×10^{-2}	1.83×10^{-2}
85°	1.96×10^{-2}	6.21×10^{-3}	2.82×10^{-3}	2.13×10^{-3}
135°	-1.04×10^{-2}	-3.17×10^{-3}	-3.21×10^{-3}	-1.41×10^{-3}
170°	3.89×10^{-4}	1.48×10^{-4}	6.68×10^{-5}	5.97×10^{-5}
ι	3.5	4.0	4.5	5
10°	1.10×10^{-2}	9.46×10^{-3}	6.56×10^{-3}	7.43×10^{-3}
45°	1.73×10^{-2}	1.52×10^{-2}	1.08×10^{-2}	7.83×10^{-3}
85°	2.66×10^{-2}	3.64×10^{-3}	6.47×10^{-4}	3.48×10^{-3}
135°	-1.12×10^{-3}	-8.46×10^{-4}	-8.82×10^{-4}	-5.59×10^{-4}
170°	8.09×10^{-5}	9.55×10^{-5}	3.06×10^{-5}	1.66×10^{-5}

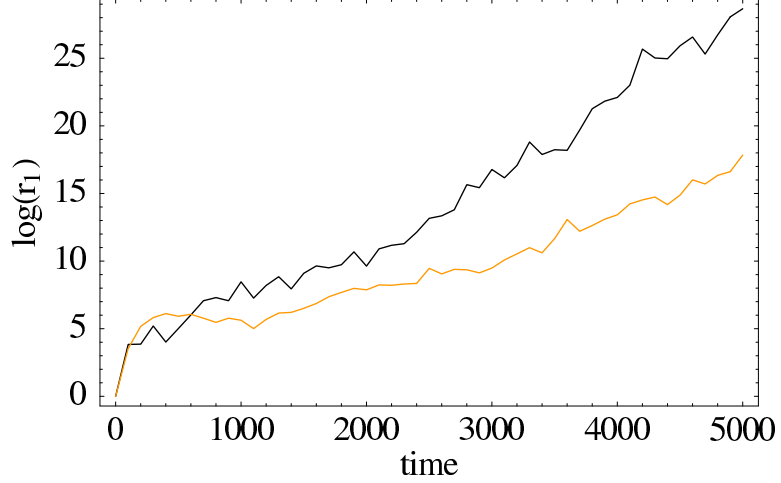


Figure 2.15: The natural logarithm of the largest ellipsoid axis vs. time for proper time τ (dark) and coordinate time t (light). The exponents are clearly different, but the Lyapunov timescales $\tau_\lambda = 1/\lambda_\tau$ and $t_\lambda = 1/\lambda_t$ are related by Eq. (2.56).

final plunge is $N \sim \frac{M}{\mu}$, which is 10^5 for the systems in Tables 2.3 and 2.4. Since the table represents values of $\Delta\phi$ for 2000 times the average radial orbital period, this means that the total phase shift during the inspiral is $50 \Delta\phi_{\text{table}}$. For a 10° inclination angle the total phase shift is on the order of a tenth of a radian to a radian. Slightly more realistic values of the number of orbits can be obtained using Fig. 2 in [36], which gives $N \sim 2 \times 10^4$ orbital periods from $r = 4M$ to the plunge at $r \approx M$ for $a = 0.998$, $\iota = 10^\circ$, and $M/\mu = 10^5$. Since the orbit spends most of its time between $4r_{\text{ms}}$ and $2r_{\text{ms}}$, interpolating in Table 2.4 gives $\Delta\phi_{\text{total}} \approx 10 \times \Delta\phi_{r=3.0} = 2 \times 10^{-2}$. This is only a rough estimate, since the orbits in [36] are circular, while the orbits we consider are eccentric.

2.7.4 Comments on time, rescaling, and norms

In this chapter, we have elected to use τ as the time parameter, a rescaling time T of $100 M$, and a projected norm (Sec. 2.4.1). Here we discuss the effects of varying these choices.

First, we consider the effects of using coordinate time t in place of τ . In Fig. 2.15, we plot the natural logarithm of the largest ellipsoid axis $\log[r_1(\tau)]$ vs. τ together with $\log[r_1(t)]$ vs. t . (We use the unrescaled deviation vector approach for simplicity, since the Jacobian approach requires a new Jacobian matrix for each coordinate change.) The exponents are $\lambda_\tau = 5.05 \times 10^{-3} M^{-1}$ and $\lambda_t = 2.51 \times 10^{-3} M^{-1}$, implying Lyapunov timescales of $\tau_\lambda = 1.98 \times 10^2 M$ and $t_\lambda = 3.98 \times 10^2 M$. The average value of $dt/d\tau$ over the orbit is 2.06, whereas $t_\lambda/\tau_\lambda = 2.01$, so the relationship

$$\frac{t_\lambda}{\tau_\lambda} = \left\langle \frac{dt}{d\tau} \right\rangle \quad (2.56)$$

discussed in the introduction is well satisfied.

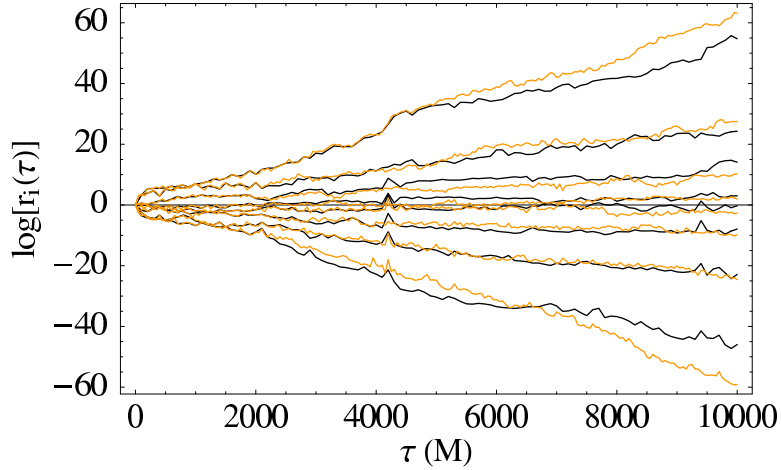


Figure 2.16: Natural logarithms of the ellipsoid axes vs. τ for rescaling time $T = 100 M$ (dark) and time $T = 50 M$ (light).

Second, we discuss the effects of varying the rescaling time T . We find that choosing T to be a moderate fraction of the shortest Lyapunov timescale (corresponding to the largest Lyapunov exponent) works best, giving each axis enough time to grow before rescaling while still keeping the negative exponents from underflowing and preventing the largest axis from dominating. Rescaling times between $50 M$ and $100 M$ work best for the systems we consider, which have Lyapunov timescales ranging from $10^2 M$ to $10^3 M$. A comparison of results for $T = 50 M$ and $T = 100 M$ appears in Fig. 2.16.

Third, we compare the projected norm used here to a naïve Euclidean norm for determining the length of the phase-space tangent vectors ξ_i . As shown in Fig. 2.17, even using a 12-dimensional Euclidean norm changes the resulting exponent very little (approximately 15% in this example). Given its conceptual advantages, we choose to use the projected norm with the confidence that the Lyapunov exponent order of magnitude is robust.

2.8 Conclusions

A spinning test particle, as described by the Papapetrou equations, appears to be chaotic in Kerr spacetime, with maximum e -folding timescales of a few $\times 10^2 M$. The applicability of this result is limited by three main factors: (1) chaos appears only for physically unrealistic values of the spin parameter; (2) other effects, such as tidal coupling, may be important for some astrophysical systems, violating the pole-dipole approximation implicit in the Papapetrou equations; and (3) we neglect gravitational radiation. The third limitation is not fatal, since the radiation timescales can be long enough that chaos, if present in the conservative limit, would have time to manifest itself in the gravitational radiation of extreme mass-ratio systems.

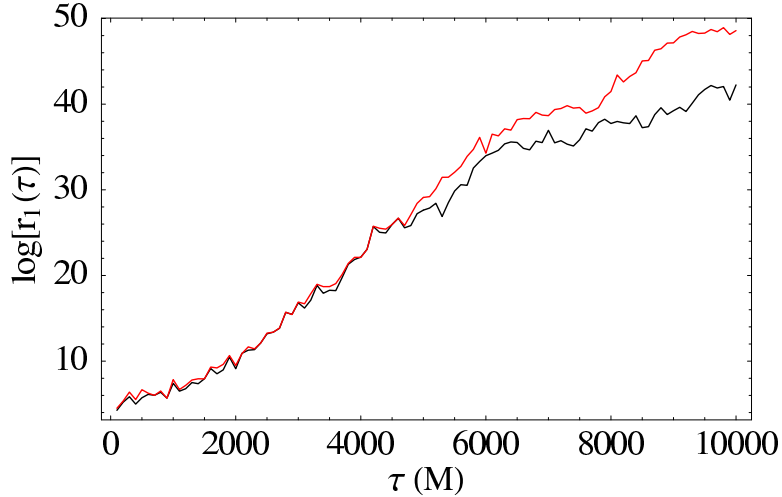


Figure 2.17: The natural logarithm of the largest ellipsoid axis vs. τ for the Euclidean norm (top) and the projected norm from Sec. 2.4.1 (bottom).

In the unphysical $S = 1$ limit, the Lyapunov exponents exhibit characteristics expected of a Hamiltonian system, appearing in pairs $\pm\lambda$ (Sec. 2.6.2). There are zero Lyapunov exponents which correspond to the constants of the motion, but the other exponents are in general nonzero. (For the Kerr orbits considered in this chapter, we find that two of the three independent exponents are nonzero, as illustrated in Fig. 2.8.) Typical orders of magnitude for the largest Lyapunov exponents are a few $\times 10^{-3} M^{-1}$ for unphysical spins ($S = 1$). For physically realistic spin parameters (Sec. 2.7.3), we find that $\lambda_{\max} \lesssim \text{few} \times 10^{-7} M^{-1}$, corresponding to e -folding timescales of a few $\times 10^6 M$. Even this bound appears to be limited only by the total integration time; in all physically realistic cases considered, λ_{\max} is indistinguishable from zero (using $S = 0$ integrations as a baseline).

From the perspective of gravitational radiation detection, our most important conclusion is that chaos seems to disappear for physically realistic values of S , i.e., values of S for which the test-particle approximation and hence the Papapetrou equations are valid. We are unable to comment on the dynamics of comparable mass-ratio binaries, since such systems are not accurately modeled by the Papapetrou equations, but for *extreme* mass-ratio binaries it appears unlikely that chaos will present a problem for the calculation of theoretical templates for use in matched filters. A more thorough exploration of parameter space is needed to reach a firmer conclusion (Chapter 3).

Acknowledgements

I would like to thank Sterl Phinney for his support, encouragement, and excellent suggestions. Thanks also to Scott Hughes for contributing through his ideas and enthusiasm for this project. I would especially like to thank Janna Levin for her careful reading of the paper on which this chapter

is based, and for insightful comments. Finally, I would like to thank Kip Thorne for teaching me general relativity and Jim Yorke for teaching me dynamical systems theory.

Appendix: Full Jacobian

For reference, we list the derivatives needed to calculate the full Jacobian matrix.

From Sec. 2.5.2, we have the following:

$$\frac{\partial \dot{x}^\mu}{\partial x^\nu} = N \left[p_\alpha g^{\alpha\mu}_{,\nu} + w^\mu_{,\nu} + v^\mu (v_\alpha w^\alpha_{,\nu} + \frac{1}{2} N w^\alpha w^\beta g_{\alpha\beta,\nu}) \right] \quad (2.57)$$

$$\frac{\partial \dot{x}^\mu}{\partial p_\nu} = N (g^{\mu\nu} + W^{\mu\nu} + N v^\mu w^\nu) + N v^\mu v_\alpha W^{\alpha\nu} \quad (2.58)$$

with

$$W^{\mu\nu} = -{}^* R^{*\mu\alpha\nu\beta} S_\alpha S_\beta \quad (2.59)$$

$$\frac{\partial \dot{x}^\mu}{\partial S_\nu} = N V^{\mu\nu} + N v^\mu v_\alpha V^{\alpha\nu} \quad (2.60)$$

with

$$V^{\mu\nu} = -S_\alpha p_\beta ({}^* R^{*\mu\alpha\beta\nu} - {}^* R^{*\mu\nu\alpha\beta}). \quad (2.61)$$

Now we simply apply the product rule many times:

$$\frac{\partial \dot{p}_\mu}{\partial x^\nu} = -p_\alpha S_\beta (R^{*\alpha\beta}_{\mu\gamma,\nu} v^\gamma + R^{*\alpha\beta}_{\mu\gamma} v^\gamma_{,\nu}) + p_\alpha (\Gamma^\alpha_{\beta\mu,\nu} v^\beta + \Gamma^\alpha_{\beta\mu} v^\beta_{,\nu}) \quad (2.62)$$

$$\frac{\partial \dot{p}_\mu}{\partial p_\nu} = -S_\beta \left(R^{*\nu\beta}_{\mu\gamma} v^\gamma + R^{*\alpha\beta}_{\mu\gamma} p_\alpha \frac{\partial v^\gamma}{\partial p_\nu} \right) + \Gamma^\nu_{\beta\mu} v^\beta + \Gamma^\alpha_{\beta\mu} p_\alpha \frac{\partial v^\beta}{\partial p_\nu} \quad (2.63)$$

$$\frac{\partial \dot{p}_\mu}{\partial S_\nu} = -R^{*\alpha\nu}_{\mu\gamma} v^\gamma p_\alpha - R^{*\alpha\beta}_{\mu\gamma} \frac{\partial v^\gamma}{\partial S_\nu} p_\alpha S_\beta + \Gamma^\alpha_{\beta\mu} p_\alpha \frac{\partial v^\beta}{\partial S_\nu}$$

$$\frac{\partial \dot{S}_\mu}{\partial x^\nu} = -p_\mu S_\alpha p_\gamma S_\delta \left(R^{*\alpha\gamma\delta}_{\beta,\nu} v^\beta + R^{*\alpha\gamma\delta}_{\beta} v^\beta_{,\nu} \right) + S_\alpha (\Gamma^\alpha_{\beta\mu,\nu} v^\beta + \Gamma^\alpha_{\beta\mu} v^\beta_{,\nu}) \quad (2.64)$$

$$\frac{\partial \dot{S}_\mu}{\partial p_\nu} = -S_\alpha S_\delta v^\beta \left(\delta_\mu^\nu R^{*\alpha\gamma\delta}_{\beta} p_\gamma + p_\mu R^{*\alpha\gamma\delta}_{\beta,\nu} \right) - p_\mu R^{*\alpha\gamma\delta}_{\beta} S_\alpha \frac{\partial v^\beta}{\partial p_\nu} p_\gamma S_\delta + \Gamma^\alpha_{\beta\mu} \frac{\partial v^\beta}{\partial p_\nu} S_\alpha \quad (2.65)$$

$$\frac{\partial \dot{S}_\mu}{\partial S_\nu} = -p_\mu p_\gamma v^\beta \left(R^{*\nu}_{\beta}{}^{\gamma\delta} S_\delta + R^{*\alpha}_{\beta}{}^{\gamma\nu} S_\alpha \right) - p_\mu R^{*\alpha}_{\beta}{}^{\gamma\delta} S_\alpha \frac{\partial v^\beta}{\partial S_\nu} p_\gamma S_\delta + \Gamma^\nu_{\beta\mu} v^\beta + \Gamma^\alpha_{\beta\mu} \frac{\partial v^\beta}{\partial S_\nu} S_\alpha \quad (2.66)$$

Accidentally leaving off the final term in $\frac{\partial \dot{S}_\mu}{\partial S_\nu}$ led to the robust but spurious chaotic behavior mentioned in Sec. 2.5.1.

Bibliography

- [1] J. Levin, Phys. Rev. Lett. **84**, 3515 (2000).
- [2] J. Levin, Phys. Rev. D **67**, 044013 (2003).
- [3] S. Suzuki and K. Maeda, Phys. Rev. D **55**, 4848 (1997).
- [4] J. D. Schnittman and F. A. Rasio, Phys. Rev. Lett. **87**, 121101 (2001).
- [5] N. J. Cornish, Phys. Rev. D **64**, 084011 (2001).
- [6] N. J. Cornish and J. Levin, [gr-qc/0207016](#).
- [7] N. J. Cornish and J. Levin, Phys. Rev. Lett. **89**, 179001 (2002).
- [8] S. A. Hughes, Class. Quantum Grav. **18**, 4067 (2001).
- [9] <http://lisa.jpl.nasa.gov>.
- [10] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems* (Springer, New York, 1997).
- [11] N. Cornish and J. Levin, Phys. Rev. Lett. **78**, 998 (1997).
- [12] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1993).
- [13] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
- [14] A. Papapetrou, Proc. R. Soc. London **A209**, 248 (1951).
- [15] W. G. Dixon, Proc. R. Soc. London **A314**, 499 (1970).
- [16] C. Møller, *Theory of Relativity* (Oxford University Press, London, 1972).
- [17] B. M. Barker and R. F. O'Connell, Gen. Relativ. Gravit. **5**, 539 (1974).
- [18] L. E. Kidder, Phys. Rev. D **52**, 821 (1995).
- [19] O. Semerák, Mon. Not. R. Astron. Soc. **308**, 863 (1999).

- [20] G. B. Cook, S. L. Shapiro, and S. A. Teukolsky, *Astrophys. J.* **424**, 823 (1994).
- [21] D. Lai, F. A. Rasio, and S. L. Shapiro, *Astrophys. J. Suppl. Ser.* **88**, 205 (1993).
- [22] M. Nauenberg, *Astrophys. J.* **175**, 417 (1972).
- [23] V. S. Geroyannis and P. J. Papasotiriou, *Astrophys. J.* **534**, 359 (2000).
- [24] P. Wiggins and D. Lai, *Astrophys. J.* **532**, 530 (2000).
- [25] T. Tanaka, Y. Mino, M. Sasaki, and M. Shibata, *Phys. Rev. D* **54**, 3762 (1996).
- [26] J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57**, 617 (1985).
- [27] V. Karas and D. Vokrouhlický, *Gen. Relativ. and Grav.* **24**, 729 (1992).
- [28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, England, 1992).
- [29] J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, NJ, 1987).
- [30] S. Hojman, Ph.D. thesis, Princeton University (1975).
- [31] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York, 1964).
- [32] A. J. Hanson and T. Regge, *Ann. Phys. (N.Y.)* **87**, 498 (1974).
- [33] S. Hughes, http://www.tapir.caltech.edu/listwg1/EMRI/hughes_geodesic.html.
- [34] L. S. Finn and K. S. Thorne, *Phys. Rev. D* **62**, 124021 (2000).
- [35] S. A. Hughes, *Phys. Rev. D* **61**, 084004 (2000).
- [36] S. A. Hughes, *Phys. Rev. D* **64**, 064004 (2001).
- [37] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).

Chapter 3

A survey of spinning test particle orbits in Kerr spacetime

This chapter has been accepted for publication in Phys. Rev. D; available as M. D. Hartl, gr-qc/0302103.

Abstract

We investigate the dynamics of the Papapetrou equations in Kerr spacetime. These equations provide a model for the motion of a relativistic spinning test particle orbiting a rotating (Kerr) black hole. We perform a thorough parameter space search for signs of chaotic dynamics by calculating the Lyapunov exponents for a large variety of initial conditions. We find that the Papapetrou equations admit many chaotic solutions, with the strongest chaos occurring in the case of eccentric orbits with pericenters close to the limit of stability against plunge into a maximally spinning Kerr black hole. Despite the presence of these chaotic solutions, we show that physically realistic solutions to the Papapetrou equations are not chaotic; in all cases, the chaotic solutions either do not correspond to realistic astrophysical systems, or involve a breakdown of the test-particle approximation leading to the Papapetrou equations (or both). As a result, the gravitational radiation from bodies spiraling into much more massive black holes (as detectable, for example, by LISA, the Laser Interferometer Space Antenna) should not exhibit any signs of chaos.

3.1 Introduction

The present chapter continues the investigation initiated in Chapter 2, which considered the dynamics of spinning test particles (i.e., compact astrophysical objects) orbiting a spinning black hole (Kerr spacetime). The primary objective is to determine whether or not the orbits of spinning compact objects spiraling into much more massive black holes are chaotic. Evidence for chaotic orbits in

relativistic binaries has been presented in [1–6] (although only [1] was directly concerned with the extreme mass ratio systems we consider here). An astrophysical example of the systems we consider is a solar-mass black hole or neutron star orbiting a supermassive black hole in a galactic nucleus. Such systems are potentially important sources of gravitational radiation in frequency bands detectable by space-based gravitational wave detectors such as the proposed Laser Interferometer Space Antenna (LISA) mission. The methods of data analysis for signals from such detectors typically rely on matched filters, which use a discrete mesh in parameter space to achieve a given signal-to-noise ratio. The presence of chaos would cause the number of templates needed to fit a given signal to grow exponentially with decreasing mesh size, potentially overwhelming the computers performing the analysis.

Many studies have used the Papapetrou equations to investigate the dynamics of spinning test particles in the background spacetime of a central black hole (including most recently [1, 7–10]; see [9] and [10] for a more thorough literature review). We found in Chapter 2 that the Papapetrou equations formally admit chaotic solutions in Kerr spacetime, but the chaos seemed to disappear for physically realistic spins. This conclusion was tentative, since we investigated only a few values of the many parameters appearing in the equations. In the present study, we undertake a thorough search of parameter space in order to determine the prevalence of chaotic orbits, both for dynamically interesting (but physically unrealistic) orbits with high spin parameter S and for astrophysically relevant systems with smaller spin.

As in Chapter 2, we use Lyapunov exponents to measure the divergence of nearby trajectories and to provide an estimate for the timescale of the divergence. For chaotic systems, initial conditions separated by a small distance ϵ_0 diverge exponentially, with the separation growing as $\epsilon(\tau) = \epsilon_0 e^{\lambda\tau}$, where λ is the Lyapunov exponent. For nonchaotic orbits, $\lambda = 0$, but when λ is nonzero it provides the e -folding timescale $\tau_\lambda = 1/\lambda$ for chaotic behavior to manifest itself. In this chapter we use two rigorous methods for determining the maximum Lyapunov exponent λ_{\max} described in Chapter 2. Further details appear below in Sec. 3.4 and in Sec. III of Chapter 2.

Our parameter space search makes use of a convenient orbital parameterization technique, which allows us to specify desired values of pericenter r_p and orbital inclination ι (Sec. 3.3). We vary these orbital parameters along with the spin magnitude, eccentricity, Kerr spin angular momentum a , and spin inclination, and calculate the largest Lyapunov exponent for each set of initial conditions. A complete description of our methods for searching parameter space appears in Sec. 3.5. Although not exhaustive, the resulting survey of parameter space gives a thorough picture of chaos in the Papapetrou equations with a Kerr background.

We set $G = c = 1$, use sign conventions as in MTW [11], and use Boyer-Lindquist coordinates (r, θ, ϕ) for Kerr spacetime. We use vector arrows for 4-vectors and boldface for Euclidean vectors. The symbol \log denotes the natural logarithm.

3.2 Spinning test particles

We use the Papapetrou equations of motion [12] to model a spinning test particle. As reformulated by Dixon [13], these equations describe the motion of a “pole-dipole particle,” which neglects effects smaller than those due to the mass monopole and spin dipole (thus neglecting the tidal coupling, which is a mass quadrupole effect). The Papapetrou equations describe a particle with negligible mass compared to the masses responsible for generating the background spacetime. In our case, we write μ for the test particle’s mass and M for the mass of the central Kerr black hole; the test particle limit then requires that $\mu \ll M$.

3.2.1 Equations of motion

Dixon’s formulation for the equations of motion uses an antisymmetric spin tensor $S^{\mu\nu}$ to represent the spin angular momentum of the particle. The covariant derivative of the 4-momentum p^μ deviates from the nonspinning (geodesic) case by a term representing a coupling of the spin to the background spacetime curvature. If we write v^μ for the 4-velocity, the full equations of motion appear as follows:

$$\begin{aligned}\frac{dx^\mu}{d\tau} &= v^\mu \\ \nabla_{\bar{v}} p^\mu &= -\frac{1}{2} R^\mu{}_{\nu\alpha\beta} v^\nu S^{\alpha\beta} \\ \nabla_{\bar{v}} S^{\mu\nu} &= 2p^{[\mu} v^{\nu]}.\end{aligned}\tag{3.1}$$

Here $R^\mu{}_{\nu\alpha\beta}$ is the Riemann curvature tensor of the spacetime, which in our case corresponds to the Riemann tensor for the Kerr metric. Note that in the case of vanishing spin (all components of $S^{\mu\nu}$ equal to zero) or flat spacetime (all components of $R^\mu{}_{\nu\alpha\beta}$ equal to zero), we recover the geodesic equation $\nabla_{\bar{v}} p^\mu = 0$.

For numerical and conceptual purposes, we use a reformulation of the equations of motion in terms of the momentum 1-form p_μ and spin 1-form S_μ . The spin 1-form can be derived from the 4-momentum and spin tensor using

$$S_\mu = -\frac{1}{2} \epsilon_{\mu\nu\alpha\beta} u^\nu S^{\alpha\beta},\tag{3.2}$$

where $u^\nu = p^\nu/\mu$. The spin 1-form is easier to visualize than the spin tensor, and the Papapetrou equations are better-behaved numerically in the $S \rightarrow 0$ limit when expressed in terms of the spin 1-form. Details of this formulation appear in Chapter 2 and the appendix.

3.2.2 General constraints

The Papapetrou system of equations (3.1) is highly constrained, even in arbitrary spacetimes. The 4-momentum and spin satisfy normalization conditions that are conserved by the equations of motion:

$$p^\mu p_\mu = -\mu^2 \quad (3.3)$$

and

$$\frac{1}{2} S^{\mu\nu} S_{\mu\nu} = S^2. \quad (3.4)$$

A further condition is required to identify a unique worldline for the center of mass:

$$p_\mu S^{\mu\nu} = 0. \quad (3.5)$$

A more detailed discussion of this “spin supplementary condition” appears in Chapter 2.

3.2.3 Kerr constraints

The Papapetrou equations share an important property with geodesic motion, namely, to each symmetry of the background spacetime (typically represented by a Killing vector) there corresponds a constant of the motion. Kerr spacetime has two such symmetries, which provide two constraints in addition to those described in the previous section. (The Killing tensor that leads to the Carter constant in the geodesic does not correspond to a conserved quantity when the spin is nonzero; see Sec. 3.3.2.3 below.)

To each Killing vector $\vec{\xi}$ there corresponds a constant given by the standard expression for geodesics, plus a contribution due to a coupling with the spin tensor:

$$C_\xi = \xi^\mu p_\mu - \frac{1}{2} \xi_{\mu;\nu} S^{\mu\nu}. \quad (3.6)$$

Kerr spacetime’s two Killing vectors $\vec{\xi}^t = \partial/\partial t$ and $\vec{\xi}^\phi = \partial/\partial \phi$ then give energy and z angular momentum conservation:

$$-E = p_t - \frac{1}{2} g_{t\mu,\nu} S^{\mu\nu} \quad (3.7)$$

and

$$J_z = p_\phi - \frac{1}{2} g_{\phi\mu,\nu} S^{\mu\nu}. \quad (3.8)$$

3.2.4 The spin parameter S

The spin magnitude S that appears in Eq. (3.4) quantifies the size of the spin and thus plays a crucial role in determining the behavior of spinning test particle systems. As in Chapter 2, we measure

distances and times in terms of M and momenta in terms of μ . In these units, S is measured in units of μM . When measured in traditional geometrized units, the spin of a black hole can be as large as its mass squared, i.e., $S_{\text{geom, max}} = \mu^2$. In this case the spin parameter goes like $S \sim \mu^2/(\mu M) = \mu/M$, which is necessarily small for a test particle system. This result applies to all astrophysically relevant systems, as shown in Chapter 2. Thus, we have the important physical constraint that *the spin parameter S must satisfy $S \ll 1$ for all physically realistic systems*. A thorough discussion of various possibilities (including neutron stars and white dwarfs) in Chapter 2 showed that realistic values of S for LISA sources fall in the range 10^{-4} – 10^{-7} .

The smallness of the spin's effect is not dependent on our choice to measure S in terms of μM . If instead we measure S in terms of μ^2 , then the equation of motion for $S^{\mu\nu}$ becomes (rewriting $\nabla_{\bar{v}}$ as $D/d\tau$)

$$\frac{D(S^{\mu\nu}/\mu^2)}{d(\tau/M)} = 2 \frac{1}{\mu} p^{[\mu} v^{\nu]}, \quad (3.9)$$

where we have factored out from each variable its corresponding scale factor. This gives

$$\frac{DS^{\mu\nu}}{d\tau} = \frac{\mu}{M} \left(2p^{[\mu} v^{\nu]} \right), \quad (3.10)$$

so that in these units the spin's effective magnitude is suppressed by a factor of μ/M . Measuring S in terms of μM effectively absorbs the unavoidable smallness of the spin effect into the spin parameter itself.

3.3 Parameterizing initial conditions

The many constraints on the equations of motion lead to considerable complexity in parameterizing the initial conditions. We summarize here the primary parameterization method described in Chapter 2, and then discuss in detail a method for parameterizing initial conditions using the geometrical properties of the orbit.

3.3.1 Energy and angular momentum parameterization

It is easy to parameterize the Papapetrou equations directly in terms of the momentum and spin components, but more physically relevant parameterizations are more difficult. The parameterization method discussed in Chapter 2 solves for the initial conditions using the integrals of the motion E and J_z . This method also serves as the foundation for a more sophisticated parameterization in terms of orbital parameters (Sec. 3.3.2).

The energy and angular momentum parameterization method proceeds as follows. If we think of the system in terms of the spin 1-form S_μ , we have 12 variables, four each for position, 4-momentum, and spin. Since Kerr spacetime is static and axially symmetric, without loss of generality we may set

the initial time and axial angle to zero: $\tau = \phi = 0$. (We use the proper time τ as our time parameter, as discussed in Sec. 3.4 below.) We then provide the initial values for r , θ , and p_r , together with the constants E , J_z , and S . Finally, we provide two components of the spin vector in an orthonormal basis. It is easiest to specify the radial and θ components $S^{\hat{r}}$ and $S^{\hat{\theta}}$; since the r - θ part of the Kerr metric is diagonal, we may then easily derive the 1-form components S_r and S_θ .

Having specified values for seven of the variables, we must now use the five equations that relate them. Since we measure the particle's momentum in terms of its rest mass μ , the momentum normalization relation is

$$p^\mu p_\mu = -1, \quad (3.11)$$

which we use to eliminate p_θ . We then solve the spin normalization condition

$$S^\mu S_\mu = S^2 \quad (3.12)$$

to eliminate S_ϕ . The spin-momentum orthogonality constraint and the two integrals of the motion then give three equations in the three remaining unknowns p_t , p_ϕ , and S_t :

$$0 = p_\mu S_\nu g^{\mu\nu} \quad (3.13)$$

$$E = -p_t + \frac{1}{2} g_{t\mu,\nu} S^{\mu\nu} \quad (3.14)$$

$$J_z = p_\phi - \frac{1}{2} g_{\phi\mu,\nu} S^{\mu\nu} \quad (3.15)$$

(The first of these equations is equivalent to the condition $p_\mu S^{\mu\nu} = 0$, as discussed in the appendix.)

We solve Eqs. (3.13)–(3.15) with a Newton-Raphson root finder, which works robustly given good initial guesses. The terms $\frac{1}{2} g_{t\mu,\nu} S^{\mu\nu}$ and $\frac{1}{2} g_{\phi\mu,\nu} S^{\mu\nu}$ are typically small, even in the physically unrealistic case of $S \sim 1$, so E and J_z are good initial guesses for their corresponding momenta. We typically use S_θ as the initial guess for S_t . In all cases, we verify *a posteriori* that all constraints are satisfied to within machine precision.

3.3.2 Orbital geometry parameterization

While the method described above is sufficient for determining a set of valid initial conditions, parameterizing orbits by energy and angular momentum is not particularly intuitive. It is much more natural to think in terms of the orbital geometry, so we prefer to parameterize by the pericenter r_p , the eccentricity e , and the orbital inclination angle ι . Such parameters have precise meaning only for geodesic orbits, but are nevertheless still useful for the case of spinning test particles.

3.3.2.1 Kerr geodesics

The first step in parameterizing solutions of the Papapetrou-Dixon equations using orbital parameters is to solve the geodesic case. The traditional method for specifying a geodesic in terms of conserved quantities uses the energy E , the z angular momentum L_z , and the Carter constant Q [11]. In order to use the orbital parameters, we adopt a mapping from (r_p, e, ι) to (E, L_z, Q) based on unpublished notes supplied by Teviet Creighton and Scott Hughes (and implemented in Hughes's Kerr geodesic integrator [14]).

In order to use the more intuitive orbital parameters, we must determine the set (E, L_z, Q) given the set (r_p, e, ι) . We obtain two of the necessary equations by noting that the radial velocity $dr/d\tau$ vanishes at pericenter and apocenter, since these radii correspond to turning points in the radial motion. The equation for the time-evolution of the Boyer-Lindquist radius r is [11]:

$$\Sigma^2 \left(\frac{dr}{d\tau} \right)^2 = R(r), \quad (3.16)$$

where

$$R(r) = [E(r^2 + a^2) - aL_z]^2 - \Delta[r^2 + (L_z - aE)^2 + Q], \quad (3.17)$$

and we use the standard auxiliary variables

$$\Sigma = r^2 + a^2 \cos^2 \theta \quad (3.18)$$

and

$$\Delta = r^2 - 2Mr + a^2. \quad (3.19)$$

The quantity a is the Kerr spin parameter J/M , i.e., the central black hole's spin angular momentum per unit mass, which is dimensionless in our normalized units. From Eq. (3.17) we see that $dr/d\tau = 0$ implies that $R(r) = 0$, so we obtain one equation at each turning point:¹

$$R(r_p) = 0 \quad (3.20)$$

and

$$R(r_a) = 0, \quad (3.21)$$

where the apocenter is defined by

$$r_a = \left(\frac{1+e}{1-e} \right) r_p. \quad (3.22)$$

¹In this chapter we never consider exactly circular orbits, but we note that our prescription fails in this case: the conditions (3.20) and (3.21) are identical when $e = 0$, since $r_a = r_p$. For exactly circular orbits one must use the additional condition $R'(r_p) = 0$, where $R' = dR/dr$.

The final equation required to complete the mapping is [15]

$$Q = L_z^2 \tan^2 \iota. \quad (3.23)$$

The value of ι resulting from this definition agrees closely with the maximum value of $|\pi/2 - \theta|$ for a numerically determined solution to the equations of motion, i.e., it faithfully captures a geometric property of the orbit.

Eqs. (3.20)–(3.23) give three equations in three unknowns, which are easy to solve using a non-linear root finder as long as good initial guesses for the energy, angular momentum, and Carter constant can be found. The approach we adopt uses the degenerate cases of circular equatorial orbits to provide the raw material for analytical guesses. The energies for prograde and retrograde circular orbits in the equatorial plane are

$$E^{\text{pro}}(r) = \frac{1 - 2v^2 + \tilde{a}v^3}{\sqrt{1 - 3v^2 + 2\tilde{a}v^3}} \quad (3.24)$$

and

$$E^{\text{ret}}(r) = \frac{1 - 2v^2 - \tilde{a}v^3}{\sqrt{1 - 3v^2 - 2\tilde{a}v^3}}, \quad (3.25)$$

where we write $v \equiv \sqrt{M/r}$ and $\tilde{a} = a/M$ for notational simplicity. The initial guess for the energy is then an average of these energies, weighted using the inclination angle,² with “radius” given by the semimajor axis of an ellipse with pericenter r_p and eccentricity e :

$$E^{\text{guess}} = \frac{1}{2} [\alpha_+ E^{\text{pro}}(r_{\text{semi}}) + \alpha_- E^{\text{ret}}(r_{\text{semi}})], \quad (3.26)$$

where

$$\alpha_{\pm} = 1 \pm \cos \iota, \quad (3.27)$$

and

$$r_{\text{semi}} = \frac{r_p}{1 - e}. \quad (3.28)$$

It is possible (though rare) for Eq. (3.26) to yield an energy guess greater than 1; in this case, we simply set $E^{\text{guess}} = 1$.

Once we have a guess for the energy, we can find the corresponding guess for the angular momentum. Using the value from Eq. (3.26) and the expression for the angular momentum for a circular equatorial orbit gives

$$L_z^{\text{guess}} = \cos \iota \sqrt{\frac{1 - e^2}{2(1 - E^{\text{guess}})}} \quad (3.29)$$

²We adopt the convention that $a \geq 0$, so that ι indicates whether the orbit is prograde ($0 \leq \iota < \frac{\pi}{2}$) or retrograde ($\frac{\pi}{2} < \iota \leq \pi$).

as an initial guess for the angular momentum. Finally, the guess for the Carter constant is

$$Q^{\text{guess}} = (L_z^{\text{guess}})^2 \tan^2 \iota. \quad (3.30)$$

Plugging Eqs. (3.26), (3.29), and (3.30) into the nonlinear root finder yields the actual values of E , L_z , and Q to within machine precision in fewer than 10 iterations.

One caveat about our parameterization method is worth mentioning: some values of (r_p, e, ι) correspond to unstable Kerr orbits, and in this case the method returns a set (E, L_z, Q) corresponding to an orbit with a different pericenter from the one requested. We can illustrate this behavior by factoring Eq. (3.17), which is a quartic function in r :

$$R(r) = (1 - E^2)(r_1 - r)(r - r_2)(r - r_3)(r - r_4). \quad (3.31)$$

The roots are ordered so that $r_1 \geq r_2 \geq r_3 \geq r_4$. Bound motion occurs for $r_1 \geq r \geq r_2$, which implies that $r_1 = r_a$ and $r_2 = r_p$, but this works only for stable orbits. In the case that the orbit *requested* is unstable, the set (r_p, e, ι) returned by the algorithm instead corresponds to a nearby, stable orbit. In this case, the numerically calculated roots of $R(r)$ satisfy $r_3 = r_{p, \text{requested}}$, but the parameterization method returns r_2 as the pericenter, i.e., it returns a nearby stable orbit with pericenter $r_p = r_2 > r_{p, \text{requested}}$. As a result, the *actual* pericenter is larger than the value requested.

We use routines from [14] to identify the boundary between stable and unstable orbits (so that the latter may be excluded), but the code has a few minor bugs, and the identification procedure is not infallible. As a result, some such orbits appear in the results below (Sec. 3.5) and can be identified by having pericenters different from those requested. It is essential to understand, however, that the orbits returned by the parameterization algorithm are never unstable, and represent perfectly valid (stable) solutions to the equations of motion.

3.3.2.2 From Kerr geodesics to Papapetrou initial conditions

Once we have the set (E, L_z, Q) for the geodesic case, we can use the well-known properties of the Kerr metric to solve for all the parameters necessary for the method described in Sec. 3.3.1 above. The first step is to force the constants of the motion to agree by setting $E_{\text{Papapetrou}} = E_{\text{Kerr}}$ and $J_z = L_z$. Next, we set $\theta_0 = \pi/2$ (corresponding to the equatorial plane), since this is the only angle guaranteed to be shared by all Kerr geodesic orbits.³ Finally, we must choose an initial value r_0 for the Boyer-Lindquist radius, which in the case of $\theta_0 = \pi/2$ coincides with the crossing of the

³This restricts our sample to Papapetrou orbits that cross the equatorial plane. This almost certainly represents the vast majority of valid Papapetrou solutions, but there remains the intriguing possibility of spinning particle solutions that orbit permanently above or below the equatorial plane. We leave an examination of this possibility to future investigators.

equatorial plane. One possibility is simply to use the average value of the pericenter and apocenter,

$$r_0 = \frac{1}{2}(r_p + r_a). \quad (3.32)$$

This radius is guaranteed to lie between pericenter and apocenter, and the prescription in Eq. (3.32) works fine for the mildly eccentric orbits considered in this chapter, but a more robust method must take into account that highly eccentric orbits should in general cross the equatorial plane near pericenter. (Requiring a plane-crossing far from pericenter would force the inclination angle to be small, which is a constraint we do not wish to impose.) This suggests choosing an initial value of r close to pericenter. One flexible method is to choose

$$r_0 = r_p(1 + \alpha e), \quad (3.33)$$

where α is a number of order unity. This reduces correctly to r_p in the $e = 0$ (circular) limit, and selects an equatorial plane crossing near pericenter in the $e \sim 1$ limit. If we set $\alpha = 2$ in Eq. (3.33), then this method agrees exactly with Eq. (3.32) in the cases $e = 0$ and $e = 0.5$, and differs from Eq. (3.32) by less than 15% when $e = 0.6$. Most of the results in this chapter use Eq. (3.32), but Eq. (3.33) is preferable in general.

Once the initial θ and r are known, we can determine all components of the Kerr 4-momentum p_K^μ , but three of the Papapetrou momenta are determined by the constraints (Sec. 3.3.1). We are therefore free to force only one of the four components of the Papapetrou momentum p_P^μ to be the same as its Kerr counterpart. For the bulk of this chapter, we set the radial components equal ($p_K^r = p_P^r$), since it is the radial momentum that is most closely tied to the stability and boundedness of the orbits. This choice results in Papapetrou orbits with pericenters and eccentricities fairly close to the Kerr geodesic values, but with much higher orbital inclination angles (Fig. 3.7). The alternate choice of $p_K^\theta = p_P^\theta$ results in Papapetrou orbits with inclinations similar to their Kerr counterparts, but with very different pericenters (Fig. 3.8). See Sec. 3.5 for further discussion.

We determine a value for the radial Papapetrou momentum p_P^r using the Kerr geodesic parameters E_K , L_z , and Q by applying Eq. (3.17) and the equation

$$\rho^2 p^r = \mu \sqrt{R(r)}, \quad (3.34)$$

where $\rho = r^2 + a^2 \cos^2 \theta$ [11]. We then convert to p_r using $p_r = p^r g_{rr}$. Proceeding exactly as in Sec. 3.3.1, we specify two of the spin components and eliminate two variables using 4-momentum and spin normalization, and then solve numerically for p_t , p_ϕ , and S_t . The result is a set of initial conditions for the Papapetrou equations with the same energy and angular momentum as a Kerr geodesic with the desired values of r_p , e , and ι .

It sometimes happens that the Papapetrou initial conditions derived in this manner specify an orbit that is unstable against plunge into the black hole. Since there is no “effective potential” for a generic Papapetrou orbit as there is for Kerr geodesics, there is no way *a priori* to detect this instability. Plunge orbits are detected at runtime by testing for radial coordinates less than the horizon radius.⁴ Orbits that plunge are removed from consideration since by definition they cannot be chaotic.

3.3.2.3 Empirical orbital parameters

In making the transition from geodesics to solutions of the Papapetrou equations, we are able to enforce the conditions $E_{\text{Papapetrou}} = E_{\text{Kerr}}$ and $J_z = L_z$, but this is no guarantee that the Papapetrou orbit has the corresponding orbital parameters r_p and ι : the spin-coupling term $-\frac{1}{2}\xi_{\mu;\nu}S^{\mu\nu}$ [cf. Eq. (3.6)] has a potentially large effect on the empirical values of the pericenter and orbital inclination. This effect is most pronounced when we consider high spin parameter values, i.e., $S \sim 1$. In these dynamically interesting cases, the *empirical* pericenter and inclination will differ in general from the values requested in the parameterization.

The empirical Papapetrou pericenter $r_{p,P}$ is easy to find: we simply integrate the orbit with a small stepsize for a large number of periods, and then record the minimum radius achieved. In practice, this works robustly, reproducing almost exactly the requested Kerr value of $r_{p,K}$ in the limit $S \ll 1$. The only exception involves values of $r_{p,K}$ that correspond to requested unstable orbits, as discussed in Sec. 3.3.2.1 above. Each of these orbits has an empirical pericenter larger than the pericenter requested: the requested pericenter corresponds to the root r_3 in Eq. (3.31), but the empirical pericenter returned by the algorithm corresponds to the larger root r_2 .

Having found the empirical pericenter for an orbit, we must next find its empirical orbital inclination angle ι_P . In order to reproduce the definition from Eq. (3.23), we need to find an analogue of the Carter constant Q for spinning test particles. Kerr spacetime has a Killing tensor $K_{\mu\nu}$ [16] that satisfies

$$K_{\mu(\nu;\alpha)} = 0, \quad (3.35)$$

which gives rise to an extra conserved quantity in the case of geodesic motion:

$$K = K_{\mu\nu}p^\mu p^\nu. \quad (3.36)$$

This quantity is related to the traditional Carter constant Q by [11]

$$K = Q + (L_z - aE)^2. \quad (3.37)$$

⁴In practice, the most common runtime error is actually a numerical underflow in the integration stepsize as the particle approaches the horizon.

When the test particle has nonzero spin, the quantity defined by Eq. (3.37) is no longer constant, but there is an analogous expression that is conserved to linear order. Adapting a result from [7], we can write this approximately conserved quantity as

$$C = K_{\mu\nu} p^\mu p^\nu - 2p^\mu S^{\rho\sigma} (f^\nu{}_\sigma f_{\mu\rho\nu} - f_\mu{}^\nu f_{\rho\sigma\nu}), \quad (3.38)$$

where

$$f_{\mu\nu} = 2a \cos \theta e_{[\mu}^1 e_{\nu]}^0 + 2r e_{[\mu}^2 e_{\nu]}^3, \quad (3.39)$$

$$f_{\mu\nu\sigma} = 6 \left(\frac{a \sin \theta}{\sqrt{\Sigma}} e_{[\mu}^0 e_\nu^1 e_{\sigma]}^2 + \sqrt{\frac{\Delta}{\Sigma}} e_{[\mu}^1 e_\nu^2 e_{\sigma]}^3 \right), \quad (3.40)$$

and the $\{e^a{}_\mu\}$ are the standard orthonormal tetrad for the Kerr metric. The effective Carter “constant” for spinning particles is then

$$Q_{\text{eff}} = C - (J_z - aE)^2, \quad (3.41)$$

where we use the full angular momentum J_z (which includes the contribution from the spin) in place of L_z . The quantity Q_{eff} is nearly but not exactly constant, so in order to define an empirical inclination angle we find the maximum effective Q over an orbit, and then define ι_P by

$$Q_{\text{eff}, \text{max}} = J_z^2 \tan^2 \iota_P. \quad (3.42)$$

As in the geodesic case, the value of ι obtained from Eq. (3.42) agrees well with the maximum value of $|\pi/2 - \theta|$ over an orbit.⁵ When $S = 0$, Eq. (3.42) reduces to the definition of the orbital inclination for geodesics, Eq. (3.23).

3.4 Lyapunov exponents

3.4.1 The principal exponent

The Lyapunov exponents for a chaotic dynamical system quantify the chaos and give insight into its dynamics (revealing, for example, whether it is Hamiltonian or dissipative). For a dynamical system of n degrees of freedom, in general there are n Lyapunov exponents, which describe the time-evolution of an infinitesimal ball centered on an initial condition. This initial ball evolves into an ellipsoid under the action of the Jacobian matrix of the system, and the Lyapunov exponents are related to the average stretching of the ellipsoid’s principal axes. We described in Chapter 2 a general

⁵This is true only when we force the Kerr and Papapetrou values of p^r to agree (Sec. 3.3.2.2), which is the case for all orbits considered except for the initial conditions shown in Fig. 3.8. In that case we revert to the simpler method of finding the maximum value of $|\pi/2 - \theta|$ over several orbits.

method for calculating all n of the system's exponents, but in the present study we are interested only in the presence or absence of chaos in the Papapetrou system, so we need only calculate the principal Lyapunov exponent, i.e., the exponent corresponding to the direction of greatest stretching. A nonzero principal exponent indicates the presence of chaos.

When studying a differentiable dynamical system, we typically introduce a set of variables $\mathbf{y} = \{y_i\}$ to represent the system's phase space, together with the autonomous set of differential equations

$$\frac{d\mathbf{y}}{d\tau} = \mathbf{f}(\mathbf{y}) \quad (3.43)$$

which determine the dynamics. Associated with each initial condition is a solution (or *flow*). The principal Lyapunov exponent quantifies the *local* divergence of nearby initial conditions, so any fully rigorous method necessarily involves the local behavior of the system, i.e., its derivative. For a multidimensional system, this derivative map is given by the Jacobian matrix,

$$(\mathbf{Df})_{ij} = \frac{\partial f_i}{\partial x^j}. \quad (3.44)$$

It is the Jacobian map that determines the time-evolution of infinitesimally separated trajectories. If we consider a point \mathbf{y} on the flow and a nearby point $\mathbf{y} + \delta\mathbf{y}$, then we have

$$\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) = \frac{d}{d\tau}(\mathbf{y} + \delta\mathbf{y}) = \frac{d\mathbf{y}}{d\tau} + \frac{d(\delta\mathbf{y})}{d\tau}, \quad (3.45)$$

so that the separation $\delta\mathbf{y}$ satisfies

$$\frac{d(\delta\mathbf{y})}{d\tau} = \mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \frac{d\mathbf{y}}{d\tau} = \mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}). \quad (3.46)$$

Since

$$\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) = \mathbf{Df} \cdot \delta\mathbf{y} + \mathcal{O}(\|\delta\mathbf{y}\|^2), \quad (3.47)$$

we can write the time-evolution of the deviation vector as

$$\frac{d(\delta\mathbf{y})}{d\tau} = \mathbf{Df} \cdot \delta\mathbf{y} + \mathcal{O}(\|\delta\mathbf{y}\|^2). \quad (3.48)$$

If we identify the “infinitesimal” deviation $\delta\mathbf{y}$ with an element $\boldsymbol{\xi}$ of the tangent space at \mathbf{y} , we effectively take the limit as $\delta\mathbf{y} \rightarrow 0$; the equation of motion for $\boldsymbol{\xi}$ is then

$$\frac{d\boldsymbol{\xi}}{d\tau} = \mathbf{Df} \cdot \boldsymbol{\xi}. \quad (3.49)$$

This equation describes the time-evolution of the separation between nearby initial conditions in a

rigorous way. We will refer to $\boldsymbol{\xi}$ as a tangent vector, since formally it is an element of the tangent space to the phase space.

If the Jacobian \mathbf{Df} in Eq. (3.49) were some constant matrix \mathbf{A} , the solution would be the matrix exponential

$$\boldsymbol{\xi}(\tau) = \exp(\mathbf{A}\tau) \cdot \boldsymbol{\xi}_0. \quad (3.50)$$

For long times, the solution would be dominated by the largest eigenvector of \mathbf{A} , and would grow like

$$\|\boldsymbol{\xi}(\tau)\| \approx e^{\lambda_{\max}\tau}, \quad (3.51)$$

where λ_{\max} is the largest eigenvalue.⁶ (Here we have used $\|\boldsymbol{\xi}_0\| = 1$.) Turning things around, if we measured the time-evolution of $\boldsymbol{\xi}$, we could find an approximation for the exponent using

$$\lambda_{\max} = \frac{\log \|\boldsymbol{\xi}(\tau)\|}{\tau}. \quad (3.52)$$

The general case follows by considering Jacobian matrices that are time-dependent. In this case, we are unable to define a unique principal exponent valid for all times, but there is still a unique *average* exponent that describes the average stretching of the principal eigenvector. Our method is to track the evolution of a tangent vector as it evolves into the principal axis of the ellipsoid. If we use $r_e(\tau) = \|\boldsymbol{\xi}(\tau)\|$ to denote the length of the longest principal ellipsoid axis, the principal Lyapunov exponent is then defined by

$$\lambda_{\max} = \lim_{\tau \rightarrow \infty} \frac{\log [r_e(\tau)]}{\tau}. \quad (3.53)$$

We use an infinite time limit in this formal definition, but of course in practice a numerical approach relies on a finite cutoff to obtain a numerical approximation.

The *Jacobian method* for determining the largest Lyapunov exponent involves solving Eqs. (3.43) and (3.49) as a coupled system of differential equations in order to follow the time-evolution of a ball of initial conditions. One possibility is then to use Eq. (3.53) to estimate the system's largest exponent. A related technique, which provides more accurate exponents, is to sample $\log [r_e(\tau)]$ at regular time intervals, and then perform a least-squares fit on the simulation data. Since $\log [r_e(\tau)] = \lambda\tau$, the slope of the resulting line then gives an estimate for the Lyapunov exponent. This is the method we implement in practice.

A less rigorous but still useful technique, which we call the *deviation vector* approach, involves solving only Eq. (3.43), but for two initial conditions: \mathbf{y}_0 and $\mathbf{y}_0 + \delta\mathbf{y}_0$. If the solutions to these initial conditions a time τ later are \mathbf{y} and \mathbf{y}' , respectively, then the approximate principal Lyapunov

⁶The most common choice for the norm $\|\cdot\|$ is the Euclidean norm, but we use a slightly different norm in the case of the Papapetrou equations (Sec. 3.4.2 below).

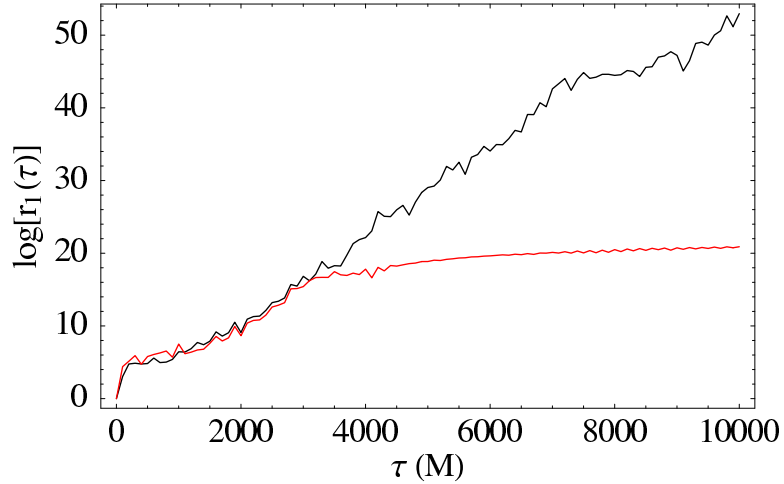


Figure 3.1: Rigorous Jacobian method compared to the unrescaled deviation vector method for an $S = 1$ particle in maximal ($a = 1$) Kerr spacetime. The vertical axis is the natural logarithm of the largest principal axis r_1 of the phase space ellipsoid; the slope is the principal Lyapunov exponent, $\lambda_{\max} \approx 5 \times 10^{-3} M^{-1}$. The unrescaled deviation vector method started with a deviation of size 10^{-7} , and it saturates at ~ 16 . This corresponds to a growth of $e^{16} \approx 9 \times 10^6$, which means that the separation has grown to a size of order unity. In conventional units, this indicates radial separations of order GM/c^2 and velocity separations of order c . The norm is calculated using the projected norm described in Sec. 3.4.2.

exponent is

$$\lambda_{\max} = \frac{\log \|\delta \mathbf{y}\| / \|\delta \mathbf{y}_0\|}{\tau}, \quad (3.54)$$

where $\delta \mathbf{y} \equiv \mathbf{y}' - \mathbf{y}$. This approach has a serious drawback: no matter how small the initial size of the deviation, eventually the method saturates as the difference between \mathbf{y} and \mathbf{y}' grows so large that it no longer samples the *local* difference between two trajectories.⁷ On the other hand, because we need only solve Eq. (3.43) and not Eq. (3.49), the deviation vector method is significantly faster than the Jacobian method (by a factor of approximately 5 for the system considered here). We therefore adopt the deviation vector method as our principal tool for broad surveys of parameter space. The method for handling the saturation problem is discussed in Sec. 3.4.3.

A comparison of the Jacobian and deviation vector methods appears in Fig. 3.1. It is apparent that the two methods agree closely until the deviation vector approach reaches the saturation limit.

3.4.2 The Papapetrou case

The discussion in the previous section was of a general nature, applicable to virtually any dynamical system described by differential equations. Here we describe some of the details needed to apply the

⁷It is possible to re-scale the deviation once it reaches a certain size, but this method is error-prone since it can depend sensitively on the precise method of rescaling. The constrained nature of the Papapetrou equations also presents difficulties for rescaling, as discussed in Sec. 3.4.2.

general methods to the Papapetrou equations. In particular, we must discuss further the key ideas of constrained deviation vectors and phase space metric.

For the Papapetrou system, the phase space vector \mathbf{y} has 12 components:

$$\mathbf{y} = (t, r, \mu, \phi, p_t, p_r, p_\theta, p_\phi, S_t, S_r, S_\mu, S_\phi). \quad (3.55)$$

The tangent vector $\boldsymbol{\xi}$ has one component for each variable. The set of equations Eq. (3.43) is simply the Papapetrou equations written out in full:

$$\begin{aligned} \dot{x}^\mu &= v^\mu \\ \dot{p}_\mu &= -R_{\mu\nu}^*{}^{\alpha\beta} v^\nu p_\alpha S_\beta + \Gamma_{\beta\mu}^\alpha p_\alpha v^\beta \\ \dot{S}_\mu &= -p_\mu \left(R^{*\alpha}{}_\beta{}^{\gamma\delta} S_\alpha v^\beta p_\gamma S_\delta \right) + \Gamma_{\beta\mu}^\alpha S_\alpha v^\beta, \end{aligned}$$

where we have used the formulation in terms of the spin 1-form described in the appendix. The second necessary equation, Eq. (3.49), requires the Jacobian matrix,

$$\begin{pmatrix} \frac{\partial \dot{x}^\mu}{\partial x^\nu} & \frac{\partial \dot{x}^\mu}{\partial p_\nu} & \frac{\partial \dot{x}^\mu}{\partial S_\nu} \\ \frac{\partial \dot{p}_\mu}{\partial x^\nu} & \frac{\partial \dot{p}_\mu}{\partial p_\nu} & \frac{\partial \dot{p}_\mu}{\partial S_\nu} \\ \frac{\partial \dot{S}_\mu}{\partial x^\nu} & \frac{\partial \dot{S}_\mu}{\partial p_\nu} & \frac{\partial \dot{S}_\mu}{\partial S_\nu} \end{pmatrix}, \quad (3.56)$$

whose explicit form appears in Chapter 2.

It is important to mention that the tangent vector $\boldsymbol{\xi}$ —or, equivalently, the deviation vector $\delta\mathbf{y}$ —cannot have completely arbitrary components. On the contrary, the deviation must be chosen carefully, in order to ensure that, given a point \mathbf{y} that satisfies the constraints from Sec. 3.2.2, the deviated vector $\mathbf{y} + \delta\mathbf{y}$ also satisfies the constraints. Otherwise, the relation $\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) = \mathbf{D}\mathbf{f} \cdot \delta\mathbf{y} + \mathcal{O}(\|\delta\mathbf{y}\|^2)$ is not satisfied. (This is the principal complication in implementing a rescaled version of the deviation vector method: the rescaled vector would violate the constraint.) In practice, we are able to find a valid deviation vector by applying the same techniques used to satisfy the constraints in the first place (Sec. 3.3); for details, see Chapter 2.

One final detail is the notion of metric: implicit in the definition of the Lyapunov exponent, Eq. (3.53), is a metric on phase space used to calculate the norm of the tangent vector $\boldsymbol{\xi}$. We adopt a metric introduced in [17], which involves projecting the deviation vector onto the spacelike hypersurface defined by the zero angular momentum observers (ZAMOs). The projection is effected using the projection tensor $P_\nu^\mu = \delta_\nu^\mu + U^\mu U_\nu$, where U^μ is the ZAMO 4-velocity. The spatial and momentum variables are then projected according to $x^\mu \rightarrow \tilde{x}^i = P_\mu^i x^\mu$, $p_\mu \rightarrow \tilde{p}_i = P_i^\mu p_\mu$, and

$S_\mu \rightarrow \tilde{S}_i = P^\mu_i S_\mu$. After the projection, we calculate the Euclidean norm in the three-dimensional hypersurface. We note that while this prescription is convenient, and reduces correctly in the nonrelativistic limit, the magnitudes of the Lyapunov exponents are similar for several other possible choices of metric Chapter 2.

3.4.3 Chaos detector

Since we are concerned with calculating Lyapunov exponents for a large number of parameter values, we use the (unrescaled) deviation vector method because of its speed. As mentioned in Sec. 3.4.1, this method has the property of saturation (as illustrated in Fig. 3.1), which is ordinarily a problem, but here we use it to our advantage as a sensitive detector of chaos.

Our method for determining whether a particular initial condition is chaotic is to consider a nearby initial condition separated by a small vector $\delta\mathbf{y}_0$ (with norm $\|\delta\mathbf{y}_0\|$ typically of order 10^{-7} or 10^{-8}) and then integrate until the system reaches 90% of the saturation limit, defined as a separation $\delta\mathbf{y}$ with unit norm. If we write $r_e = \|\delta\mathbf{y}\|/\|\delta\mathbf{y}_0\|$, the approximate Lyapunov exponent satisfies

$$\log[r_e(\tau)] = \lambda\tau, \quad (3.57)$$

so that λ is the slope of the line $\log[r_e(\tau)]$ vs. τ . Saturation occurs when $r_e = 1/\|\delta\mathbf{y}_0\|$, so that the integration ends when $\log[r_e(\tau)] = -\log(0.9\|\delta\mathbf{y}_0\|)$. We record the value of $\log[r_e(\tau)]$ at regular time intervals (typically every time $T = 100 M$ in our case), and upon reaching 90% saturation perform a least-squares fit on the simulation data to determine the exponent.

We note that the cutoff value of 0.9 is somewhat arbitrary and is the result of numerical experimentation. When using the (unrescaled) deviation vector approach, most of the chaotic systems saturate—that is, plots of $\log[r_e(\tau)]$ vs. τ flatten out (Fig. 3.1)—when the separation is of order unity, corresponding to radial separations of order M , velocity separations of order 1, and angular separations of order 1 radian. The 90% prescription ends the integration before the growth flattens out, so that the numerical estimate for the exponent is still reasonably accurate.

For the large maps of parameter space, we typically integrate up to $\tau_{\text{final}} = 10^5 M$ or saturation, whichever comes first. We choose this maximum time mainly for practical reasons: it is the longest integration possible in a reasonable amount of time. [We integrate as deeply as $10^7 M$ (Sec. 3.5.7) for individual orbits, but such long integrations are impractical for more than a handful of parameter combinations.] Dramatically longer integrations are also not particularly useful, since the timescale for gravitational radiation reaction is on the order of $\tau_{\text{GW}} = M^2/\mu$ [18], which for the most relevant LISA sources is 10^4 – $10^6 M$ (i.e., $\mu \sim 10^{-6}$ – $10^{-4} M$). Searching for chaos in such systems on a timescale longer than $\sim 10^7 M$ is probably pointless, since the radiation reaction would dominate the dynamics in this case. Finally, it appears that chaos, when present in the Papapetrou system,

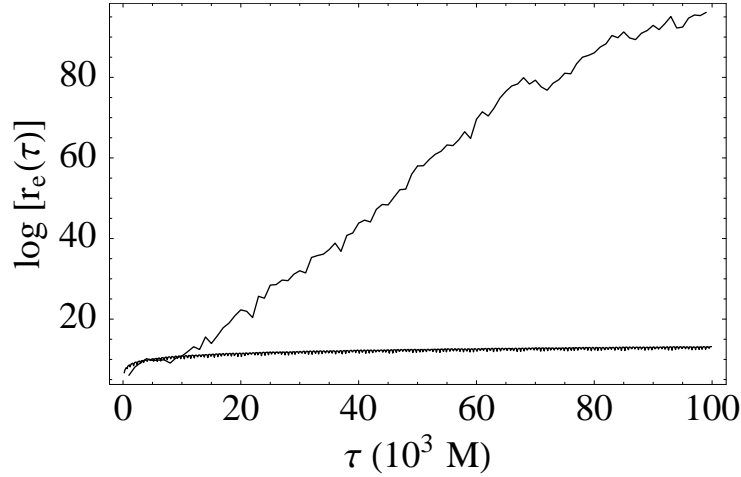


Figure 3.2: A comparison of chaotic and nonchaotic initial conditions. The slopes of the lines are the approximate Lyapunov exponents. Each initial condition shares the same values of $S = 0.1$, $e = 0.6$, and $r_p = 1.21$. They differ only in inclination angle: $\iota = 31^\circ$ (chaotic) and $\iota = 28.5^\circ$ (not chaotic). Their respective Lyapunov exponents are $\lambda = 1.0 \times 10^{-3} M^{-1}$ and $\lambda = 3.0 \times 10^{-5} M^{-1}$. See Sec. 3.5.2 for details.

manifests itself on relatively short timescales (10^2 – $10^4 M$), or else not at all. The onset of chaos is marked by a qualitative change from power law growth (which appears as logarithmic growth on our plots of $\log[r_e(\tau)]$ vs. τ) to exponential growth (which is linear on the same plots). An example of two similar initial conditions giving rise to qualitatively different dynamical behavior appears in Fig. 3.2.

One important refinement to the technique described above is to require several 90% saturation points *in a row* before declaring the orbit to be chaotic. This is necessary because some nonchaotic orbits have very high amplitudes on plots of $\log[r_e(\tau)]$ vs. τ (Fig. 3.3), even though nearby trajectories separate linearly with time (Fig. 3.4) rather than exponentially (Fig. 3.5). This phenomenon occurs mainly for orbits with many periods in the deeply relativistic zone near the horizon. Such orbits may reach “90% saturation” briefly as part of their oscillation, but quickly return to separations far below our chaotic cutoff. We therefore adopt the criterion of *three 90% saturation points in a row* (with a time $T = 100 M$ sampling interval) as a robust practical test for chaos. See Sec. 3.5.8 for further discussion.

Our confidence of this method’s robustness derives from comparing the method above to the Jacobian method for the same initial conditions. Since the Jacobian method does not saturate, the agreement of the two methods indicates that our procedure provides an accurate detector for chaos (as in Fig. 3.12 below).

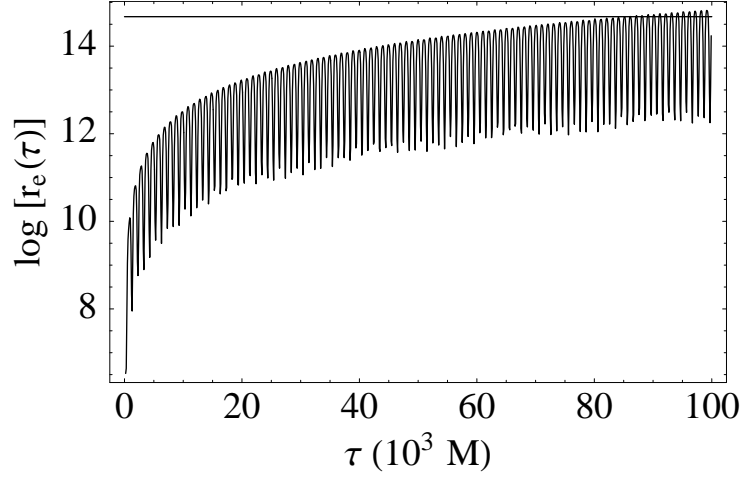


Figure 3.3: A chaos mimic: $\log[r_e(\tau)]$ vs. τ for an $S = 0$ orbit. The size of the initial deviation vector is $\epsilon_0 = 3.3 \times 10^{-7}$. The value of $\log[r_e(\tau)]$ periodically rises up to the saturation level (shown as a horizontal line at 14.82, since $e^{14.82} \epsilon_0 = 0.9$). The system's spin satisfies $S = 0$, and is hence fully integrable, which implies no chaos. We detect such spurious chaos by demanding several saturation points in a row for a positive detection. The corresponding orbit appears in Fig. 3.22.

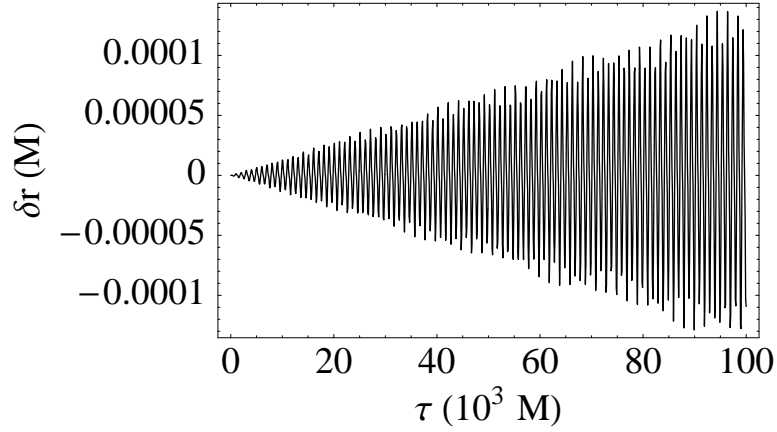


Figure 3.4: The difference δr between the Boyer-Lindquist radii of two nearby trajectories for a chaos mimic. The growth in the separation is substantial, but not exponential. The initial conditions are the same as in Fig. 3.3.

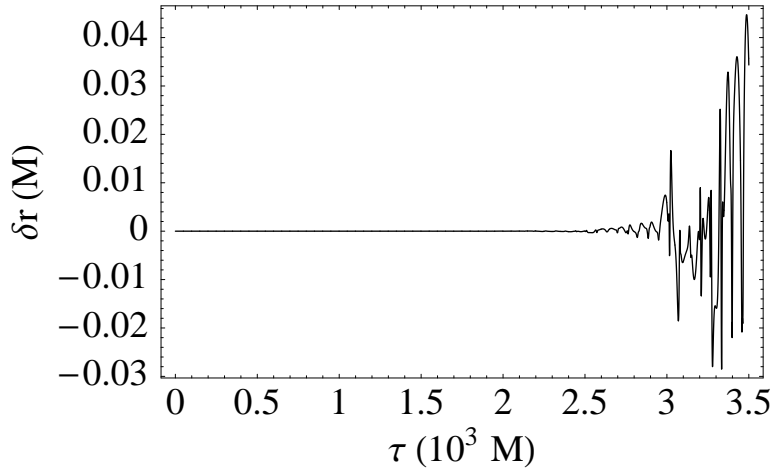


Figure 3.5: The difference δr between the Boyer-Lindquist radii of two nearby trajectories for a chaotic orbit. The initial condition is taken from the inner region of Fig. 3.9 ($r_{p,K} = 2.0$, $e_K = 0.5$, $\iota_K = 10^\circ$, $S = 1$). On this linear scale, the separation seems to burst unexpectedly, in contrast to the relatively smooth linear growth for the nonchaotic orbit shown in Fig. 3.4.

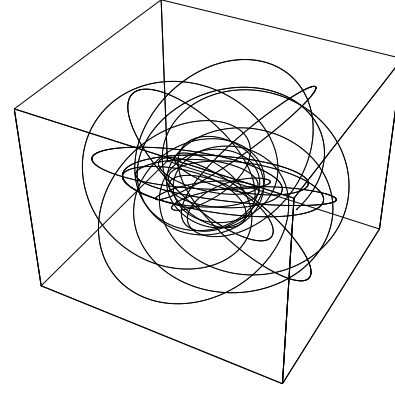
3.4.4 Implementation notes

We integrate the Papapetrou equations on a computer using Bulirsch-Stoer and Runge-Kutta integrators implemented in the C programming language, as described in Chapter 2 (and shown in Fig. 3.6). The derivatives and Jacobian matrix are extensively hand-optimized for speed. We monitor errors using constraints and conserved quantities, with a global error goal of 10^{-13} . The errors are at the 10^{-11} level or better for highly chaotic orbits after $10^5 M$. Orbits with low spin or high pericenter are even more accurate, often achieving the error goal of 10^{-13} .

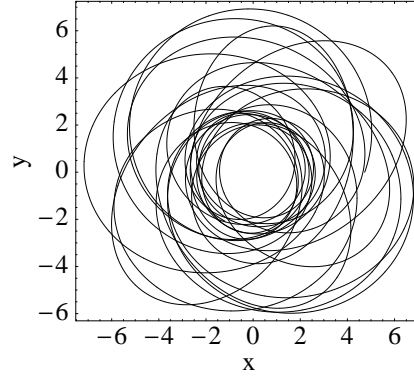
The many plots in Sec. 3.5 are typically generated using driver programs written in the Perl programming language, which in turns calls the C integrator repeatedly. This general paradigm—using an interpreted language such as Perl to call optimized routines in a compiled language such as C—is one we warmly recommend.

3.5 Results

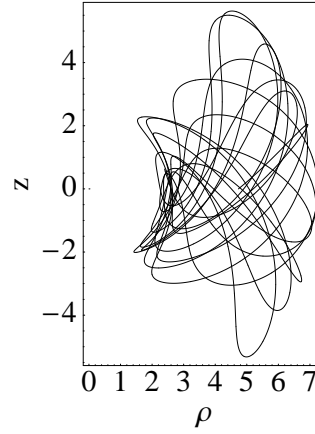
We present here the results of parameter variation in the Papapetrou system of equations (Figs. 3.7–3.43). We represent the effects compactly using several different kinds of plots, most of which involve plotting inclination vs. pericenter, with other parameters held fixed. We refer to these as $r_{p-\iota}$ maps. As discussed in Sec. 3.3, starting with the Kerr values $r_{p,K}$, ι_K , and e_K , we find the corresponding energy E_K , angular momentum $L_{z,K}$, and Carter constant Q_K , and then force the Papapetrou values to satisfy $E_P = E_K$, $J_{z,P} = L_{z,K}$ and $p_P^r = p_K^r$. Each $r_{p-\iota}$ map has two components: part (a), shown on the left, uses the Kerr parameters ι_K and $r_{p,K}$ requested by the parameterization (Sec. 3.3.2.2),



(a)



(b)



(c)

Figure 3.6: The orbit of a maximally spinning ($S = 1$) test particle in maximal ($a = 1$) Kerr spacetime, plotted in Boyer-Lindquist coordinates. (a) The orbit embedded in three-dimensional space, treating the Boyer-Lindquist coordinates as ordinary spherical polar coordinates; (b) $y = r \sin \theta \sin \phi$ vs. $x = r \sin \theta \cos \phi$; (c) z vs. $\rho = \sqrt{x^2 + y^2}$. The requested orbital inclination angle is $\iota = 6^\circ$, but the strong spin coupling gives rise to an empirical value closer to $\iota_P = 46^\circ$. The empirical pericenter is $r_p = 2.219 M$, which is fairly close to the requested value of $2.367 M$.

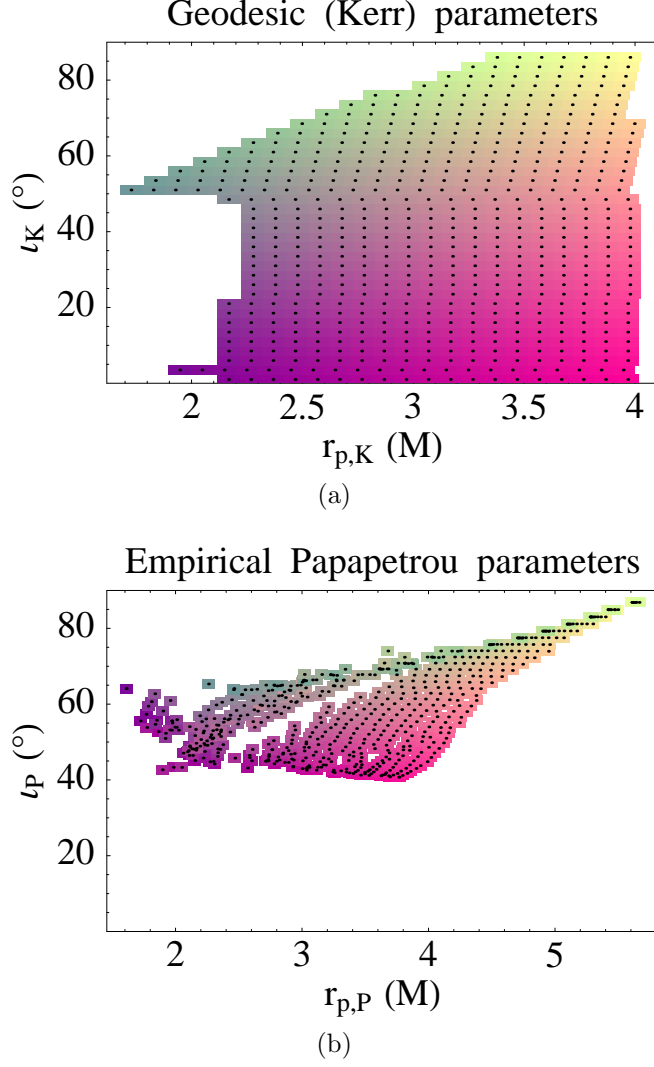


Figure 3.7: Shading/coloring of parameter space for an $S = 1$ orbit. The shading/color on the left is repeated on the right, so that we can visually determine the mapping of $(r_{p,K}, \iota_K)$ to $(r_{p,P}, \iota_P)$. It is evident that orbits with low requested pericenters and orbital inclination angles are mapped to low-pericenter orbits at high inclinations, and the entire parameter space is compressed. Note that the gap in stable initial conditions visible in (b) is a true gap, not a fold.

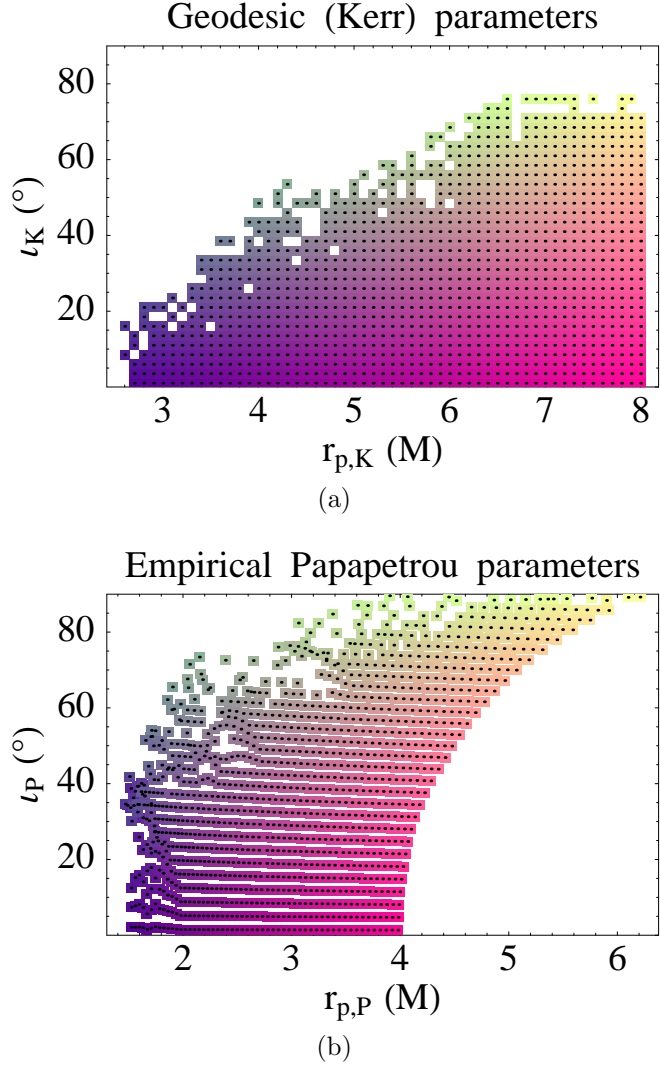


Figure 3.8: Shading/coloring of parameter space for an $S = 1$ orbit illustrating the alternate parameterization method from Sec. 3.3.2.2. As in Fig. 3.7, the shading/color on the left is repeated on the right. The spatial part of the initial spin is purely in the z direction. The Kerr p_K^θ and Papapetrou p_P^θ values are forced to agree, which leads to similar inclination angles in parts (a) and (b), at the cost of dramatically different pericenters.

while part (b), shown on the right, always shows the empirical Papapetrou values ι_P and $r_{p,P}$ in the sense of Sec. 3.3.2.3.

One important feature of r_p - ι parameter space apparent in the empirical plots is the prevalence of large empirical inclination angles for all values of the Kerr inclination ι_K . Fig. 3.7 shows the mapping for $S = 1$ between the requested Kerr parameters [Fig. 3.7(a)] and the empirical Papapetrou parameters [Fig. 3.7(b)] using a shading scheme (which appears as a more informative color scale in electronic versions of this chapter). We see that even the orbits at the bottom of Fig. 3.7(a) get mapped to high empirical inclination angles; $\iota_K = 1^\circ$ orbits are mapped to inclinations of order $\iota_P = 40^\circ$.

This compression of parameter space is the result of our choice to force the Kerr and Papapetrou value of the radial momentum to agree (Sec. 3.3.2.2). The price we pay for this choice is that the θ component of the Papapetrou momentum—which is constrained to satisfy the equations in Sec. 3.2.2—is relatively large. This high p_P^θ flings even orbits with low requested values of ι_K to high inclinations (Fig. 3.7). It is possible to find low-inclination Papapetrou orbits by requiring that the Kerr and Papapetrou values of p^θ agree, but at the cost of forcing p^r to be very different—again a result of the constraints. The resulting parameter space (Fig. 3.8) is not nearly as compressed in inclination angle, but the Papapetrou pericenters are compressed and shifted down, and many requested values of r_p , K are lost as plunge orbits. In addition, the empirical values of the eccentricity are typically not close to the value requested (reaching, e.g., $e_P = 0.75$ for $e_K = 0.5$). Because of the deficiencies of this alternate parameterization method, we choose the fixed p^r plots are our primary investigative tool in this chapter.

3.5.1 Varying pericenter and orbital inclination

In this section we show r_p - ι maps (Fig. 3.9 and Figs. 3.25–3.29) for orbits with fixed eccentricity $e = 0.5$ and spin parameters $S^{\hat{r}} = S^{\hat{z}} = 0.2S$, where S is the total spin.⁸ We indicate the strength of the chaos at each point with a grayscale rectangle, with the darkest colors representing the strongest chaos (and with white indicating no chaos). An example of such a plot appears in Fig. 3.9. (Fig. 3.10 shows a similar plot for the alternate parameterization method that forces $p_K^\theta = p_P^\theta$. The maximum exponents in the two cases are virtually identical.) The most important general results evident from the plots are two-fold. First, the largest exponents occur for orbits with pericenters deep in the relativistic zone near the horizon. Second, the prevalence of chaotic orbits is a decreasing function of spin parameter S , with virtually all chaos gone by the time $S = 0.1$.

An example of the value of the empirical $r_{p,P}$ - ι_P plots appears in Fig. 3.26, which shows orbits of particles with spin $S = 0.5$. The appearance of a strongly chaotic point in Fig. 3.26 seems

⁸These “fixed” spin parameters give rise to a variety of spin inclination angles in the fiducial (ZAMO) rest frame; see Sec. 3.5.6 below.

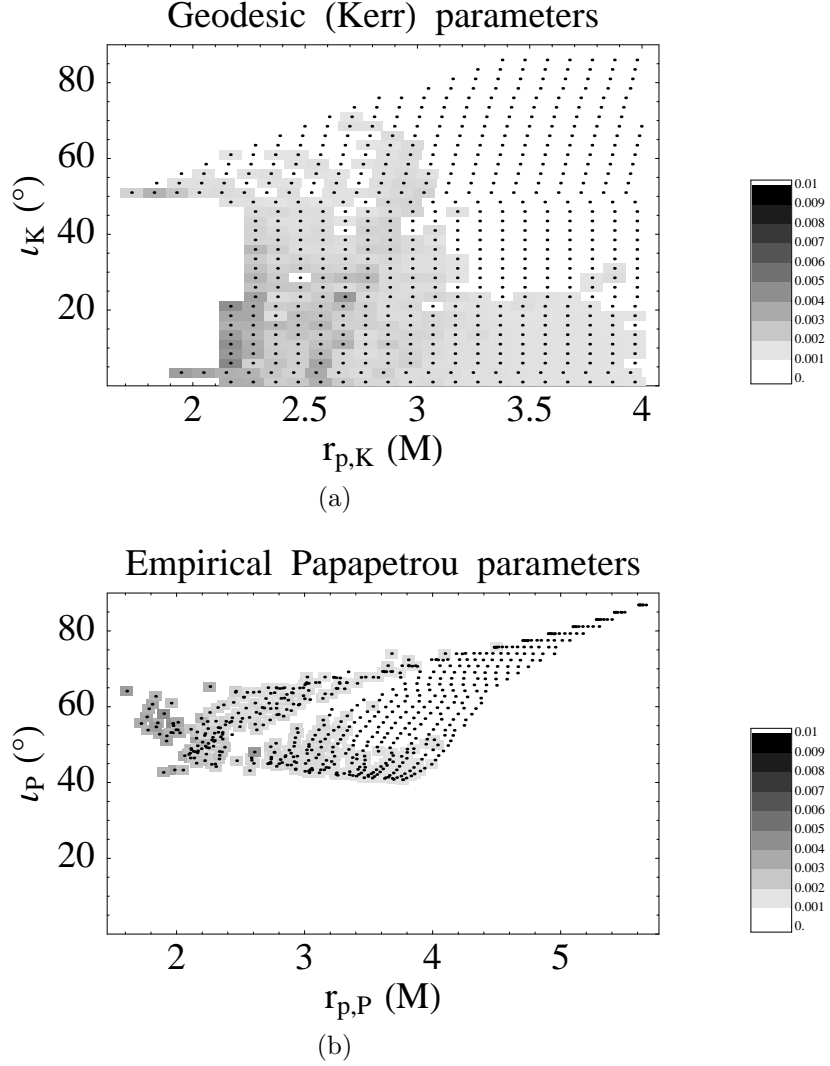


Figure 3.9: r_p - ι map for $a = 1$, $e = 0.5$, $S = 1$, $S^{\hat{r}} = S^{\hat{z}} = 0.2S$: chaos strength as a function of pericenter and orbital inclination angle. (a) Requested parameters; (b) empirical parameters. The initial conditions are the same as in Fig. 3.7. The rectangles are shaded according to the Lyapunov exponent for the initial condition represented by each point, with darker shades of gray representing larger exponents and hence stronger chaos. The point at $r_{p,K} = 2.3$, and $\iota_K = 20^\circ$ is one of the points in Fig. 3.17, which shows the effects of varying $S^{\hat{r}}$ and $S^{\hat{z}}$. The largest exponent in this plot is $\lambda = 4.1 \times 10^{-3} M^{-1}$, corresponding to a timescale of $1/\lambda = 2.4 \times 10^2 M$ for a factor of e divergence in nearby initial conditions.

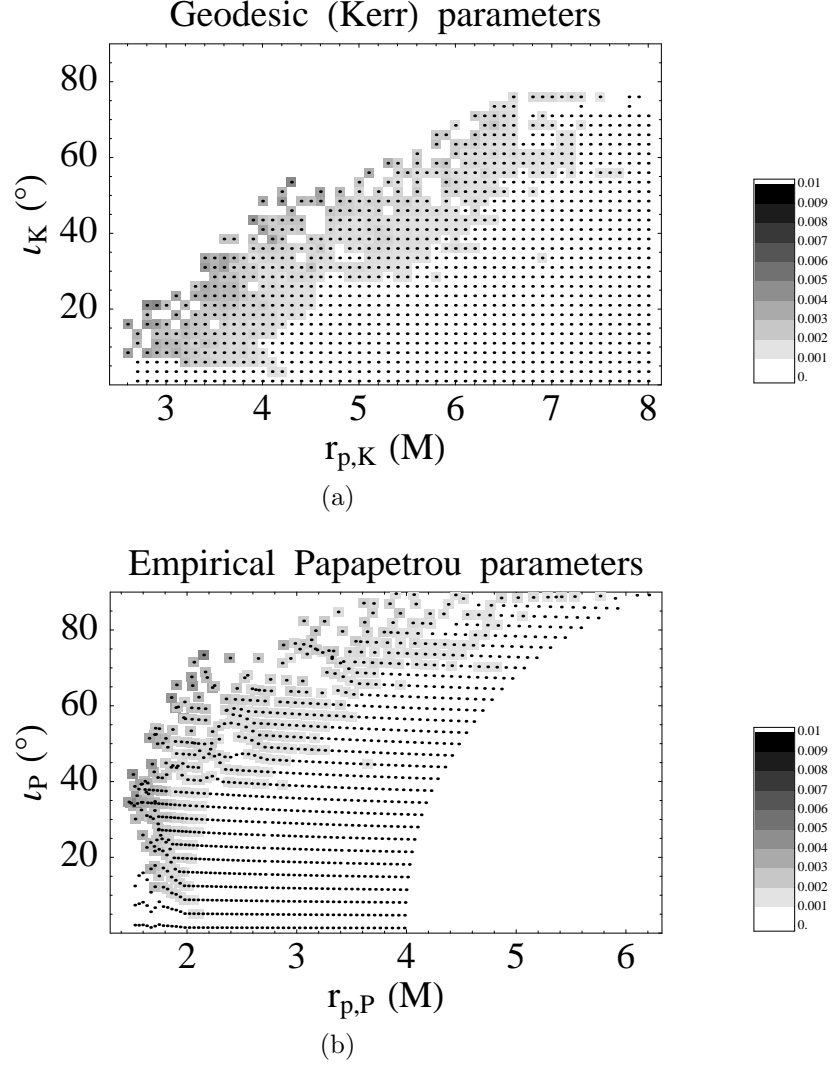


Figure 3.10: r_p - ι map for $a = 1$, $e = 0.5$, $S = 1$, using the alternate parameterization method that forces $p_K^\theta = p_P^\theta$. The initial spatial component of the spin is purely in the z direction. (a) Requested parameters; (b) empirical parameters. The initial conditions are the same as in Fig. 3.8. The rectangles are shaded according to the Lyapunov exponent for the initial condition represented by each point, with darker shades of gray representing larger exponents and hence stronger chaos. The largest exponent in this plot is $\lambda = 4.0 \times 10^{-3} M^{-1}$, corresponding to a timescale of $1/\lambda = 2.5 \times 10^2 M$ for a factor of e divergence in nearby initial conditions. These values are virtually identical to the values in Fig. 3.9.

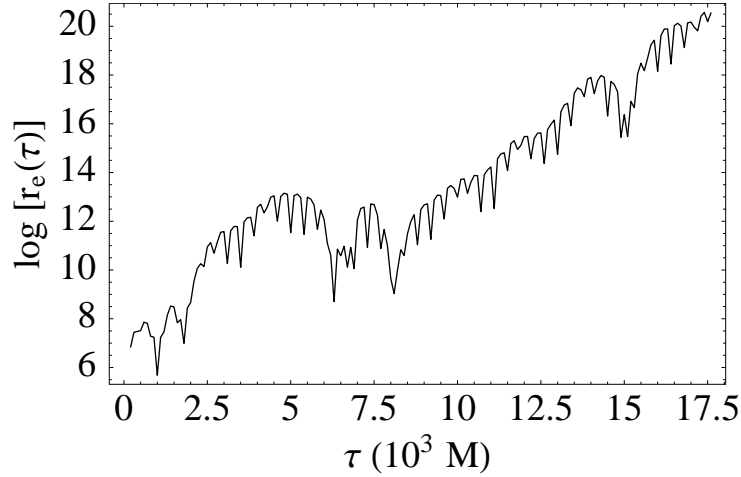


Figure 3.11: Natural logarithm of the largest ellipsoid axis vs. time. We use the unrescaled deviation vector method to investigate the single chaotic initial condition from Fig. 3.32, which has $r_{p,K} = 1.21$, $e_k = 0.6$, $\iota_K = 31^\circ$, and $S = 0.1$. The principal ellipsoid axis grows until it saturates at $\tau = 17600 M$.

perplexing, given that it is surrounded by many points with much smaller exponents. As is evident from the empirical plot, this point of strong chaos (which is, in fact, the largest Lyapunov exponent for any of the plots) maps to a compressed area of initial conditions with small empirical pericenters [Fig. 3.26(b)].

From an astrophysical standpoint, the most interesting parameter to vary is the spin S , since only $S \ll 1$ orbits are physically realistic (Sec. 3.2.4). From Figs. 3.25–3.29, which involve varying S from 0.9 down to 10^{-4} , we see that both the prevalence and strength of chaos decrease significantly as S is decreased. The Lyapunov exponent ranges as high as $10^{-2} M^{-1}$ when $S = 0.5$ (Fig. 3.26), but the chaos is weak when $S = 0.2$ (Fig. 3.27) and is gone below $S = 0.1$ (Figs. 3.28 and 3.29).

3.5.2 Varying eccentricity

The choice of $e = 0.5$ in the previous section is partially motivated by likely gravitational wave sources for LISA [19], e.g., a neutron star or small black hole in an eccentric orbit around a supermassive black hole. In this section, we consider a second series of eccentric orbits at fixed $e = 0.6$ and varying spin parameter. We also investigate the case of a near-circular orbit ($e = 0.01$) more appropriate for the circularized gravitational wave sources important for ground-based detectors such as LIGO.

The r_p - ι plots for $e = 0.6$ follow the same pattern as those with $e = 0.5$. Chaos is strongest for orbits with small pericenters and values of S of order unity (Figs. 3.30–3.33). There is a single orbit at $S = 0.1$ that appears to be chaotic (Fig. 3.32), but other than this one exception there is apparently no chaos below $S = 0.1$. A close examination of the single $S = 0.1$ chaotic orbit appears in Figs. 3.11 and 3.12, which shows that the chaos is real.

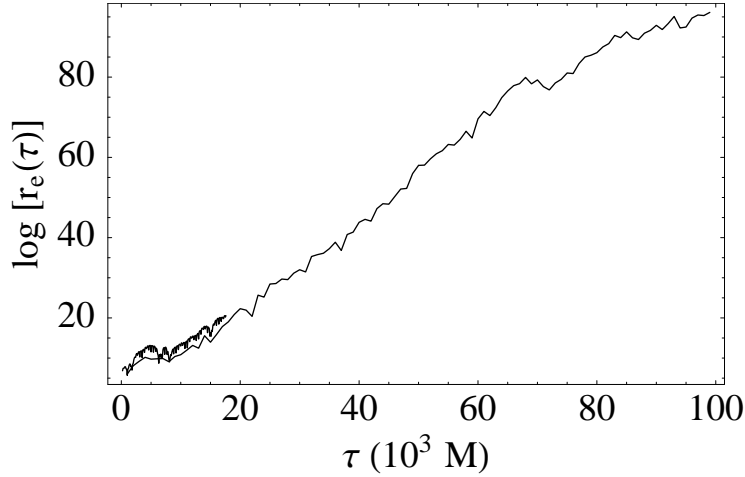


Figure 3.12: Natural logarithm of the largest ellipsoid axis vs. time using the unrescaled deviation vector method and the rigorous Jacobian method. The unrescaled integration is identical to that shown in Fig. 3.11. The two methods agree until the deviation method saturates, at which point we stop the deviated vector integration. The continued growth of the Jacobian method confirms that the orbit determined by the initial condition is indeed chaotic.

The effect of chaos is smaller for the near-circular ($e = 0.01$) orbits considered (Figs. 3.34–3.37), with typical Lyapunov exponents of order $2 \times 10^{-3} M^{-1}$ when $S = 1$ (Fig. 3.34). Moreover, we find only four points with nonzero exponents for $S = 0.5$ (Fig. 3.35), in contrast to the more eccentric orbits, which have strong chaos for $S = 0.5$. By the time $S = 0.1$ (Fig. 3.36), all chaos has completely disappeared for the near-circular orbits.

3.5.3 Varying Kerr parameter a

Here we investigate the effect of the Kerr parameter a on the strength and prevalence of chaos. Although Suzuki and Maeda found in [1] that there is chaos even in Schwarzschild spacetime, to which Kerr spacetime reduces when $a = 0$, the chaotic orbits found in [1] are exceptional orbits lying on the edge of a generalized effective potential. We found evidence in Chapter 2 that such chaotic orbits are rare.

The conclusion that chaotic orbits become less prevalent as $a \rightarrow 0$ is supported by a more thorough examination, as illustrated in Figs. 3.38–3.43. All of these orbits have eccentricity $e = 0.5$ and spin $S = 1$, and a varies from 0.9 down to 0. Even for low values of a , the parameter space is strongly affected by the spin, with plots of the empirical pericenter and orbital inclination showing significant distortion. Nevertheless, we see unambiguously that the chaotic orbits prevalent when $a = 1$ are greatly suppressed as a decreases, with no chaotic orbits at all below $a = 0.2$. This appears to be a result of the increase of the marginally stable radius r_{ms} as $a \rightarrow 0$. When $a = 0$, the minimum stable radius is at $r_{\text{ms}} = 4$ in units of the central black hole’s mass, fully $2M$ away

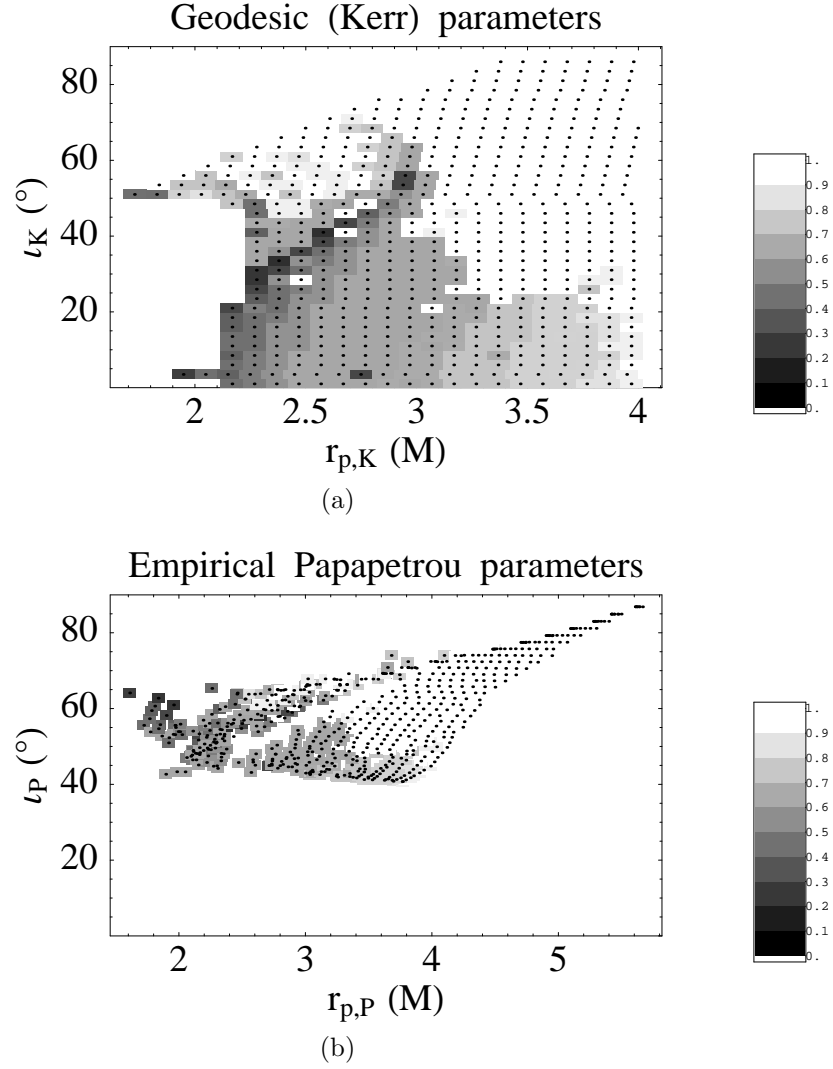


Figure 3.13: Spin cutoff map for $a = 1$ and $e = 0.5$. The rectangles are shaded according to the minimum value of S for which the corresponding initial condition is still chaotic. White rectangles indicate points that are not chaotic even when $S = 1$. The darkest points correspond to a cutoff value of $S = 0.18457$; below this critical value, none of the initial conditions are chaotic.

from the horizon at $r_H = 2M$. As discussed in Chapter 2, the extra (spherical) symmetry of the Schwarzschild metric leads to an additional integral of the motion, which also has a suppressive effect on chaos.

3.5.4 Spin cutoffs for chaos

In this section we provide *spin cutoff values* for chaos, i.e., the minimum spin values for which chaos exists. For a given initial condition defined in terms of fixed orbital parameters (as described in Sec. 3.3.2), we vary the spin parameter S and find the maximum value for which chaos occurs. The smaller this cutoff value, the stronger the chaos: nonchaotic orbits have a cutoff value of $S = 1$, i.e., they are not chaotic even in the extreme $S = 1$ limit; conversely, a cutoff value of $S = 10^{-5}$ would indicate chaos for the (physically realistic) value $S = 10^{-5}$, but not for any smaller values.⁹

Fig. 3.13 is an example of a spin cutoff map. The procedure for producing such a map is similar to the method used to make the Lyapunov r_p - ι maps: we consider a grid of points in r_p - ι space, and for each point we find an approximate value for the spin cutoff. We begin by finding out if the system is chaotic for $S = 1$, using the Lyapunov map as a start. If the orbit for $S = 1$ is chaotic, we halve the spin value and calculate the Lyapunov exponent for $S = 0.5$. If the system is still chaotic, we consider $S = 0.25$; otherwise, we consider $S = 0.75$. The procedure continues until the difference between chaotic and nonchaotic spin values is smaller than some threshold, which we choose to be 0.05. (This value is chosen to achieve reasonable accuracy while still completing the calculations in a tolerable amount of time.)

Figs. 3.13–3.15 show the spin cutoff values for several parameter combinations corresponding to Lyapunov maps from Sec. 3.5.1 above. The plots are color-coded with grayscale so that the most chaotic points—those with the *smallest* spin cutoff values—appear darkest. The points surrounded by white are not chaotic even for $S = 1$. The cutoffs for the darkest points depend on the parameter values: for Fig. 3.13 ($a = 1$ and $e = 0.5$), the same as Fig. 3.9 above), we find $S_{\text{cutoff}} = 0.18457$; Fig. 3.14 ($a = 1$ and $e = 0.01$, the same as Fig. 3.34) has $S_{\text{cutoff}} = 0.65625$; and Fig. 3.15 ($a = 0.5$ and $e = 0.5$, the same as Fig. 3.40) has $S_{\text{cutoff}} = 0.28125$. We should not take the digits of precision seriously, since these values are accurate only to within 0.05, but *in all cases the spin cutoff values are significantly above the physically realistic range of $S \sim 10^{-4}$ – 10^{-7} .*

3.5.5 Retrograde orbits

We have considered a wide variety of orbits—varying eccentricity and Kerr parameter for different pericenter, orbital inclination, and spin parameters—but so far all orbits have satisfied $0 < \iota < \pi/2$,

⁹Implicit in this scheme is the assumption of monotonicity, i.e., monotonically decreasing chaos as S decreases. While not strictly true (as discussed in Chapter 2), this assumption is still broadly applicable, and exceptions are rare.

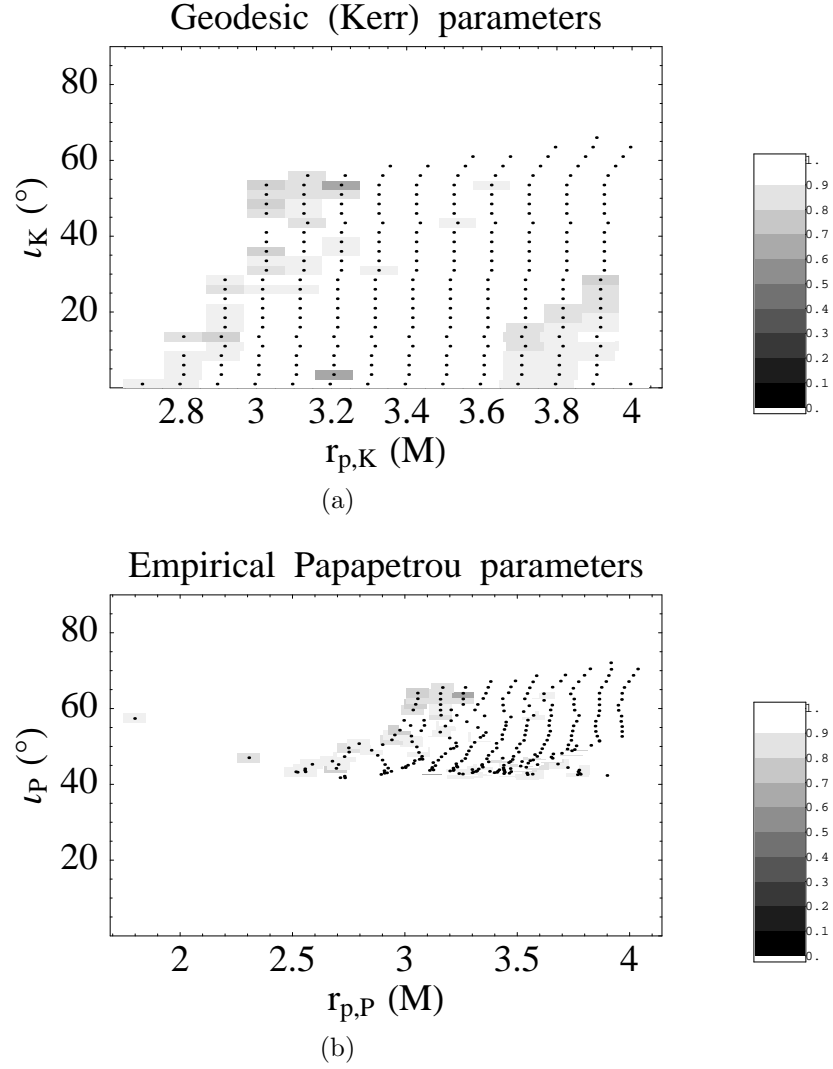


Figure 3.14: Spin cutoff map for $a = 1$ and $e = 0.01$. The rectangles are shaded according to the minimum value of S for which the corresponding initial condition is still chaotic. White rectangles indicate points that are not chaotic even when $S = 1$. The darkest points correspond to a cutoff value of $S = 0.65625$; below this critical value, none of the initial conditions are chaotic.

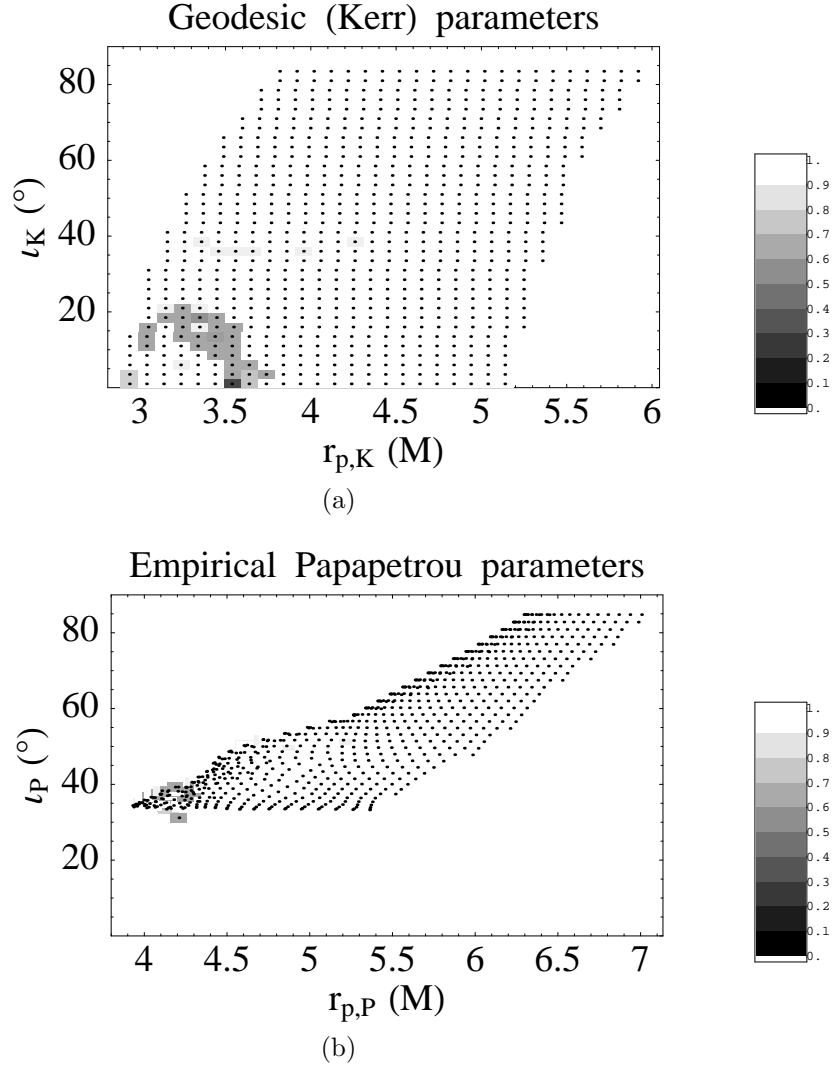


Figure 3.15: Spin cutoff map for $a = 0.5$ and $e = 0.5$. The rectangles are shaded according to the minimum value of S for which the corresponding initial condition is still chaotic. White rectangles indicate points that are not chaotic even when $S = 1$. The darkest points correspond to a cutoff value of $S = 0.28125$; below this critical value, none of the initial conditions are chaotic.

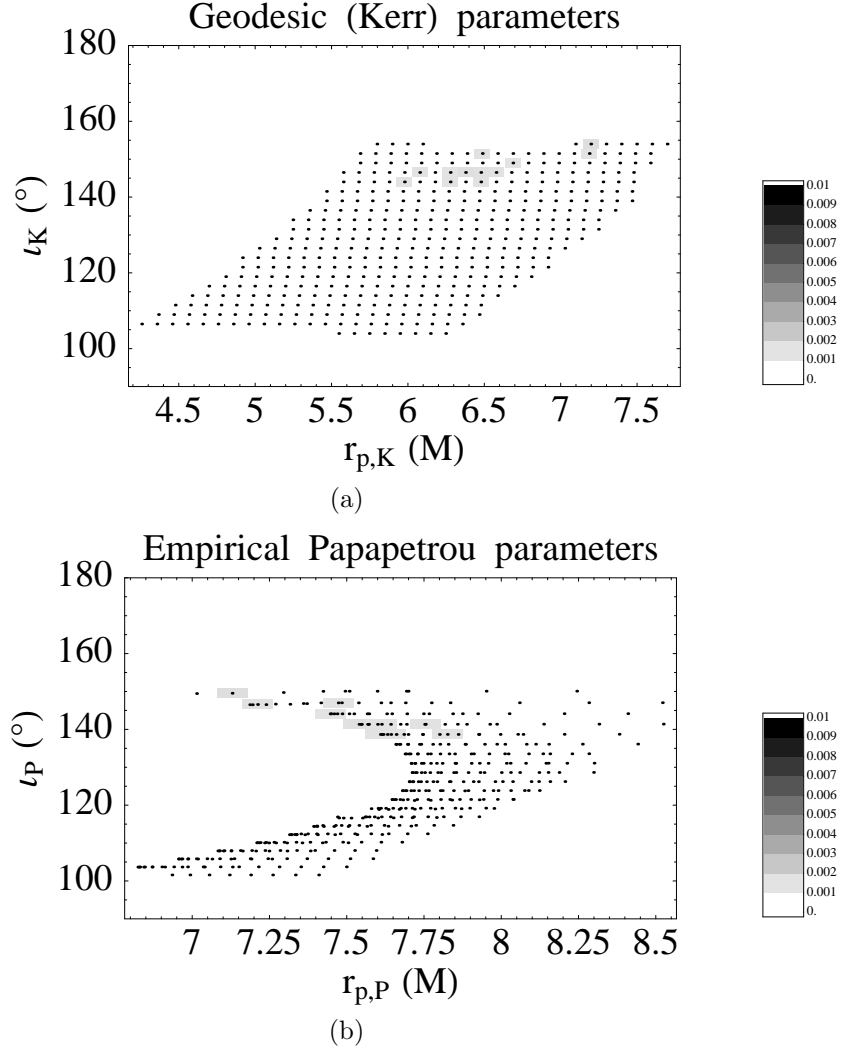


Figure 3.16: r_p - i map of retrograde orbits ($i > 90^\circ$) for $a = 1$, $e = 0.5$, $S = 1$: chaos strength as a function of pericenter and orbital inclination angle. (a) Requested parameters; (b) empirical parameters. The rectangles are shaded according to the Lyapunov exponent for the initial condition represented by each point, with darker shades of gray representing larger exponents and hence stronger chaos. The scaling is the same as in Fig. 3.9; an exponent of $\lambda = 0.01 M^{-1}$ would appear black. The chaos is weak for these retrograde orbits: the largest Lyapunov exponent in the plot is $\lambda = 3.5 \times 10^{-4} M^{-1}$.

i.e., they have all been prograde orbits, moving in the same direction as the central black hole's spin. We investigate now the case of retrograde orbits, and show that they are poor candidates for chaos.

It is evident from looking at an r_p - ι plot of a retrograde orbit (Fig. 3.16) that the pericenters are much larger than their prograde counterparts. For the $S = 1$ particle illustrated in Fig. 3.16, the minimum empirical pericenter is larger than $6M$, in contrast to prograde orbits, which get below $1.5M$. Furthermore, although it is clear from Fig. 3.16(b) that the parameter space is severely distorted, the chaos is extremely weak. The largest Lyapunov exponent, even in this extreme $S = 1$ case, is $\lambda_{\max} = 3.5 \times 10^{-4}$, two orders of magnitude smaller than in the prograde case. Unsurprisingly, all chaos disappears when $S \ll 1$. The smallest value of S_{cutoff} is 0.65265 for the parameter values shown in Fig. 3.16; we find no evidence of chaos below this value of S .

3.5.6 Varying spin inclination

So far we have always used the same values for the two spin components passed to the parameterization procedure (scaled by the total spin S): $S^{\hat{r}} = S^{\hat{z}} = 0.2S$. We consider now the effect of varying these parameters, and also discuss the corresponding initial spin inclination angles in a fiducial rest frame.

We begin with an initial condition that is chaotic for $S = 1$ but is otherwise arbitrary. We then vary $S^{\hat{r}}$ and $S^{\hat{z}}$ and calculate the Lyapunov exponent for each configuration. The result for $a = 1$, $e_K = 0.5$, $r_{p,K} = 2.3$ and $\iota_K = 20^\circ$ appears in Fig. 3.17. When $S^{\hat{r}} = S^{\hat{z}} = 0.2$, the parameter values chosen correspond to a point from Fig. 3.9. There are not valid initial conditions for all choices of parameter; in particular, negative values of $S^{\hat{z}}$ are often unstable or are unable to satisfy the spin constraints. Nevertheless, there is a large variety of parameters that do give rise to valid orbits, and many of them are chaotic. The strongest chaos exists for orbits that have small values of $S^{\hat{z}}$, but this appears to be mainly because such orbits are able to achieve small empirical pericenters. As Fig. 3.18 clearly shows, the Lyapunov exponent generally decreases as pericenter increases, with no chaotic orbits above $r_{p,P} = 2.5M$.

While we are forced by the constraints to parameterize the equations of motion in terms of spin components, it is more convenient to visualize the spin in a fiducial rest frame that is hypersurface-orthogonal to the particle's trajectory. In Kerr spacetime, the hypersurface-orthogonal observers are the zero angular momentum observers (ZAMO), which is the same frame we use when calculating the Lyapunov exponents. By projecting the components of the spin vector S^μ into the ZAMO frame in the same way as we do for the projected norm (Sec. 3.4.2), we can find the *local* value of the spin inclination angle θ_{local} (i.e., the angle between the spin axis and the axis of the central black hole). The results appear in Table 3.1. It is clear that our variation of spin components samples a large variety of initial spin inclination angles.

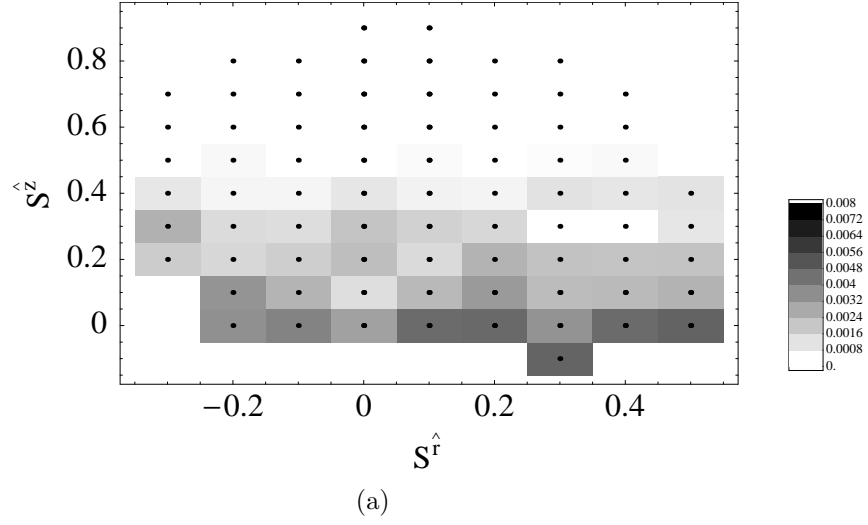


Figure 3.17: Lyapunov exponents for varying values of initial spin components $S^{\hat{r}}$ and $S^{\hat{z}}$. The parameter values $S = 1$, $r_{p,K} = 2.3$, $e_K = 0.5$, and $\iota_K = 20^\circ$ are held fixed. The point with $S^{\hat{r}} = S^{\hat{z}} = 0.2$ appears in Fig. 3.9. The actual local spin inclination angles in a fiducial (ZAMO) rest frame appear in Table 3.1. Note that only one valid initial condition exists for negative initial $S^{\hat{z}}$.

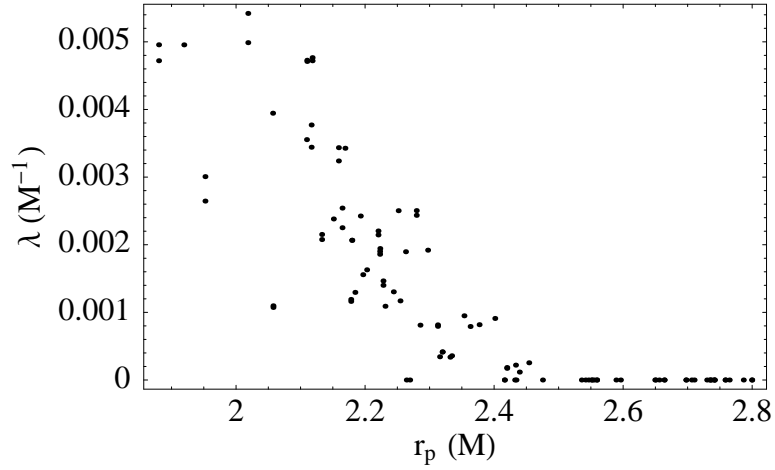


Figure 3.18: Scatter plot of empirical pericenters $r_{p,P}$ vs. Lyapunov exponent for the spin inclinations in Fig. 3.17. The Lyapunov exponent is primarily a function of pericenter, regardless of spin inclination.

Table 3.1: Local spin inclination angles θ_{local} in a fiducial (ZAMO) rest frame as a function of $S^{\hat{r}}$ and $S^{\hat{z}}$. The parameter values $S = 1$, $r_{p,K} = 2.3$, $e_K = 0.5$, and $\iota_K = 20^\circ$ are held fixed. An illustration of their Lyapunov exponents appears in Fig. 3.17.

$S^{\hat{z}}$	$S^{\hat{r}}$									
	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	
0.9				22.7°	28.7°					
0.8		48.6°	40.4°	36.1°	40.2°	48.4°	55.9°			
0.7	60.7°	54.4°	47.8°	44.7°	47.9°	54.5°	60.8°	65.7°		
0.6	64.9°	59.6°	54.2°	51.7°	54.5°	59.9°	65.2°	69.3°		
0.5	69.1°	64.6°	60.2°	58.2°	60.5°	65.0°	69.4°	72.8°		
0.4	73.2°	69.6°	66.0°	64.4°	66.3°	70.0°	73.5°	76.3°	78.4°	
0.3	77.3°	74.6°	71.8°	70.5°	72.0°	74.9°	77.6°	79.7°	81.3°	
0.2	81.5°	79.6°	77.7°	76.8°	77.9°	79.8°	81.7°	83.1°	84.2°	
0.1		84.7°	83.7°	83.3°	83.9°	84.9°	85.8°	86.5°	87.1°	
0.0		90.0°	90.0°	90.0°	90.0°	90.0°	90.0°	90.0°	90.0°	
-0.1							94.2°			

3.5.7 Deep integrations

Since we adopted a maximum time of $10^5 M$ for the Lyapunov integrations, it is reasonable to ask whether chaos might manifest itself on a longer timescale. This is certainly possible, but it appears that most initial conditions are either chaotic on a timescale of order 10^2 – $10^4 M$ or are not chaotic at all, as discussed in Sec. 3.4.3. An example appears in Fig. 3.2, where one initial condition is unambiguously chaotic, while a second located close by is not chaotic, even on a much longer timescale.

To convince ourselves that slow chaos is not lurking in the apparently nonchaotic regions, we performed a few longer-time integrations. In particular, we calculated the Lyapunov exponents using $\tau_{\text{final}} = 10^7 M$ for all the innermost ($r_p = 1.32$) orbits from Fig. 3.28 ($S = 0.1$) and Fig. 3.29 ($S = 10^{-4}$), which are strongly chaotic when $S = 1$ but are apparently without chaos below $S = 0.1$. The largest exponent occurs for $\iota = 28.5^\circ$, which is therefore the worst-case scenario. Plots of the Lyapunov exponents vs. time appear in Fig. 3.19 ($S = 0.1$) and Fig. 3.20 ($S = 10^{-4}$); their magnitudes are on the order of $\lambda_{\text{max}} = 3 \times 10^{-7} M^{-1}$, corresponding to an e -folding timescale of approximately $3.3 \times 10^6 M$. For comparison, we show λ vs. τ for a chaotic initial condition in Fig. 3.21; it is clear that the Lyapunov exponent asymptotes to a nonzero value in much less than $10^5 M$, even for weak chaos. (The initial condition in Fig. 3.21 is the $S = 0.1$ orbit illustrated in Figs. 3.11 and 3.12, whose Lyapunov exponent is actually quite small compared to analogous $S = 1$ orbits.) The long integrations thus provide strong evidence that the disappearance of nearly all chaotic orbits below $S = 0.1$ is a real effect.

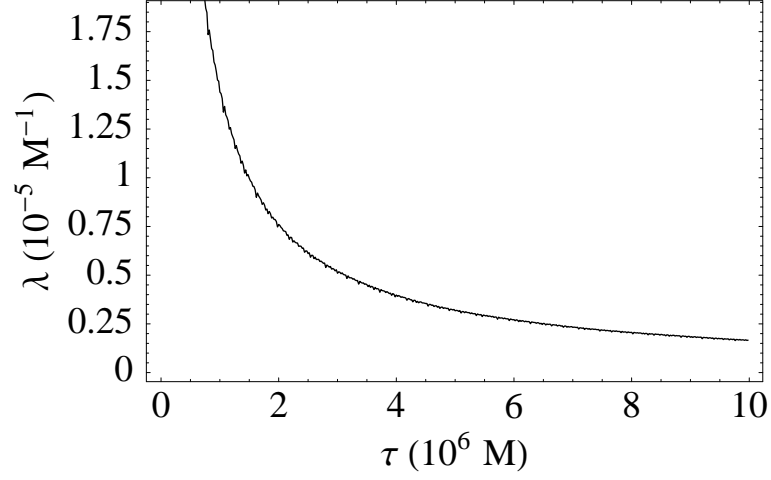


Figure 3.19: Approximate Lyapunov exponent vs. time for a nonchaotic deep ($\tau_{\text{final}} = 10^7 M$) integration. The parameter values are $a = 1$, $S = 0.1$, $e = 0.5$, $r_p = 1.32 M$, and $\iota = 28.5^\circ$, corresponding to one of the inner orbits from Fig. 3.28. The Lyapunov exponent appears to be zero; its time-evolution has the characteristic hyperbolic shape expected as $\log[r_e(\tau)]/\tau$ approaches zero for large times. A least-squares fit of $\log[r_e(\tau)]$ vs. τ gives a value of $\lambda \approx 2.8 \times 10^{-7} M^{-1}$.

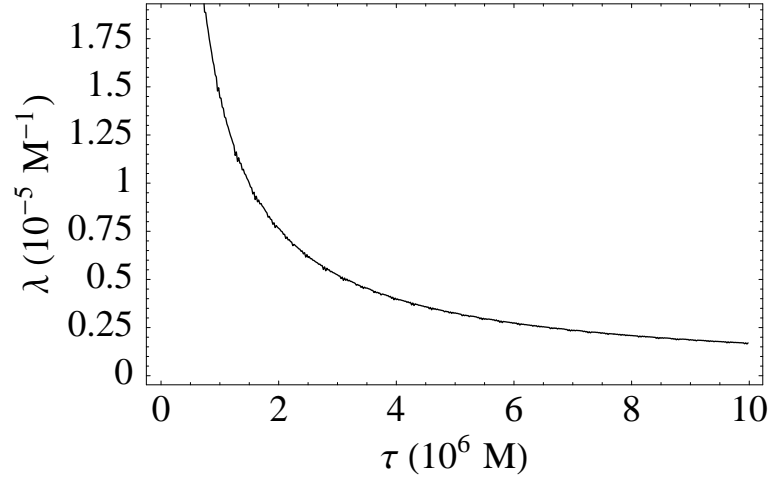


Figure 3.20: Approximate Lyapunov exponent vs. time for a nonchaotic deep ($\tau_{\text{final}} = 10^7 M$) integration. The parameter values are $a = 1$, $S = 10^{-4}$, $e = 0.5$, $r_p = 1.32 M$, and $\iota = 28.5^\circ$, corresponding to one of the inner orbits from Fig. 3.29. As in Fig. 3.19, the Lyapunov exponent appears to be zero. A least-squares fit of $\log[r_e(\tau)]$ vs. τ gives a value of $\lambda \approx 3.0 \times 10^{-7} M^{-1}$.

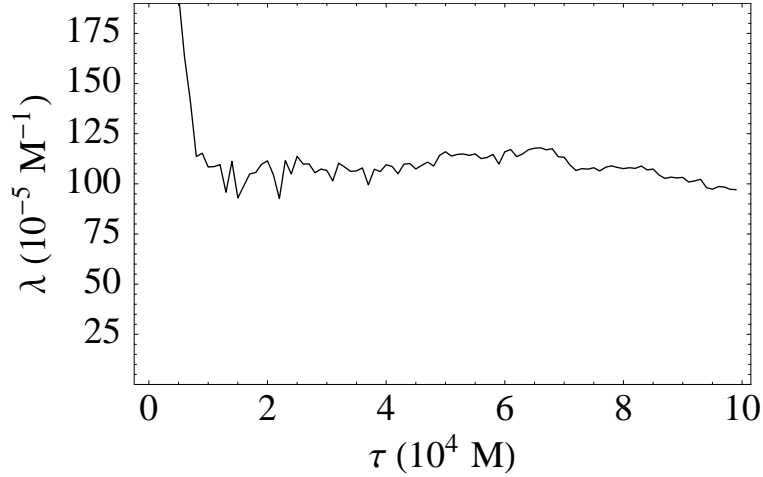


Figure 3.21: Approximate Lyapunov exponent vs. time for a chaotic deep ($\tau_{\text{final}} = 10^5 M$) integration. The simulation data is identical to that shown in Fig. 3.12; the parameter values are $a = 1$, $S = 0.1$, $e_K = 0.6$, $r_{p,K} = 1.21 M$, and $\iota_K = 31^\circ$, corresponding to the chaotic orbit from Fig. 3.32. The Lyapunov exponent for this chaotic orbit levels off after less than $10^5 M$, in contrast to Figs. 3.19 and 3.20, where λ continues to decrease even after $10^7 M$.

3.5.8 $S = 0$ and chaos mimics

Since the case of $S = 0$ corresponds exactly to geodesic orbits in Kerr spacetime, such systems cannot be chaotic—the return of the Carter constant Q and the loss of the spin degrees of freedom make the system fully integrable. Nevertheless, even some geodesic orbits can have a large separation of nearby initial conditions, which can appear to be chaotic. These chaos mimics typically spend many orbital periods whirling around deep in the strongly relativistic zone near the horizon, only occasionally zooming out to higher radii. These so-called *zoom-whirl* orbits may provide significant challenges to detection despite their formal integrability.

An example of how much divergence an $S = 0$ orbit can experience appears in Fig. 3.3. A picture of the corresponding orbit (visualized in Boyer-Lindquist coordinates embedded in ordinary space) appears in Fig. 3.22, which makes clear the large number of low-radius ϕ -periods characteristic of zoom-whirl orbits. A second example of a chaos mimic appears in Figs. 3.24 and 3.23. This orbit, in contrast to the previous one, does not have a particularly small pericenter, but its high inclination angle and zoom-whirliness allow it to mimic chaotic orbits.

The chaos mimics can exhibit large growth of the initial deviation vector, approximately a factor of 10^6 – 10^8 , after $10^5 M$. The principal means for detecting them is by requiring several high-separation points in a row (on a time- $T \approx 100 M$ basis) as mentioned in Sec. 3.4.3; the mimics have oscillations with high amplitudes due to their zoom-whirliness, but they do not represent true saturations of the separation vector.

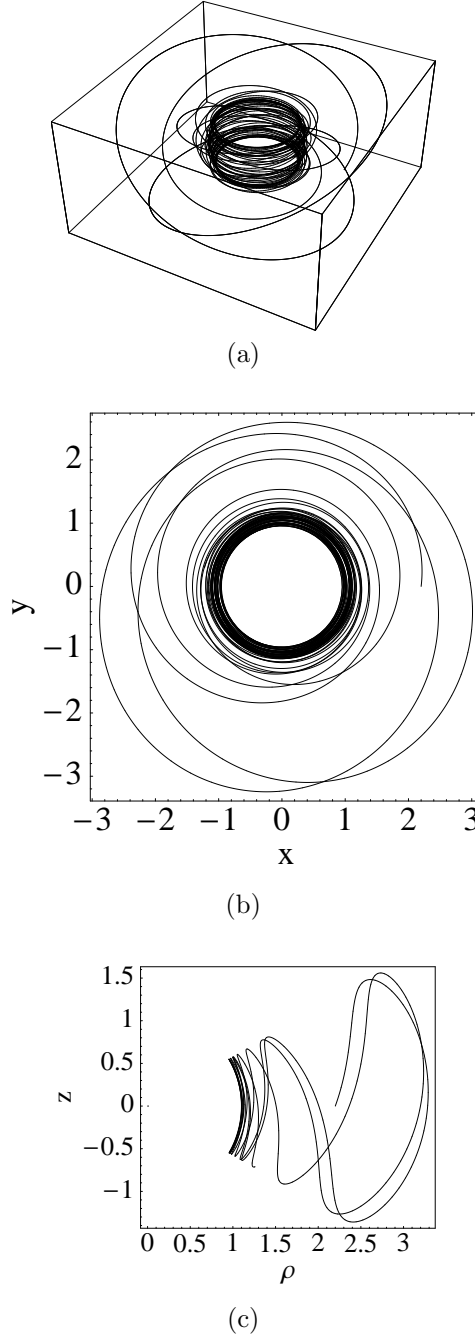


Figure 3.22: The orbit of a chaos mimic: a non-spinning ($S = 0$) test particle in maximal ($a = 1$) Kerr spacetime, plotted in Boyer-Lindquist coordinates. A Lyapunov plot of these initial conditions appears in Fig. 3.3 from Sec. 3.4.3. (a) The orbit embedded in three-dimensional space, treating the Boyer-Lindquist coordinates as ordinary spherical polar coordinates; (b) $y = r \sin \theta \sin \phi$ vs. $x = r \sin \theta \cos \phi$; (c) z vs. $\rho = \sqrt{x^2 + y^2}$. The inclination angle is $\iota = 31^\circ$, while the pericenter is $r_p = 1.1 M$ (just $0.1 M$ above the horizon at $r_H = 1 M$).

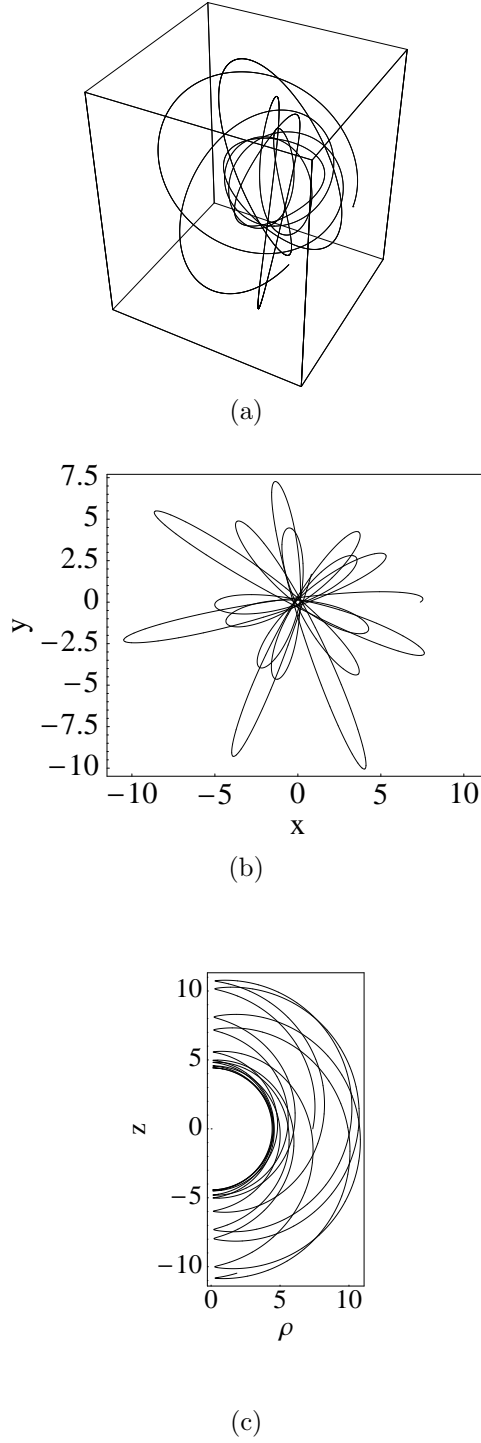


Figure 3.23: Another chaos mimic: a non-spinning ($S = 0$) test particle in maximal ($a = 1$) Kerr spacetime, plotted in Boyer-Lindquist coordinates. A Lyapunov plot of these initial conditions appears in Fig. 3.24. (a) The orbit embedded in three-dimensional space, treating the Boyer-Lindquist coordinates as ordinary spherical polar coordinates; (b) $y = r \sin \theta \sin \phi$ vs. $x = r \sin \theta \cos \phi$; (c) z vs. $\rho = \sqrt{x^2 + y^2}$. The inclination angle is $\iota = 88.5^\circ$, while the pericenter is $r_p = 4.4 M$.

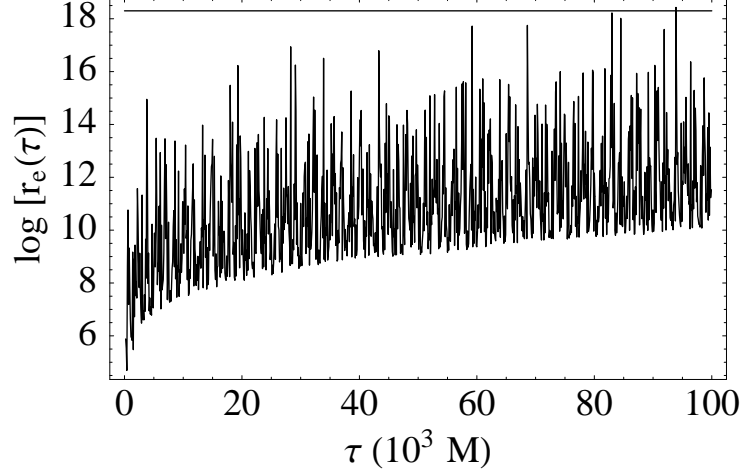


Figure 3.24: A chaos mimic: the natural logarithm of the principal ellipsoid axis ($\log[r_e(\tau)]$) vs. τ for an $S = 0$ orbit. The size of the initial deviation vector is $\epsilon_0 = 10^{-8}$. The value of $\log[r_e(\tau)]$ briefly rises up to the saturation level at $\log(0.9/\epsilon_0)$ [so that $\epsilon(\tau) = 0.9$], but the orbit is not chaotic since its spin satisfies $S = 0$. A plot of the corresponding orbit appears in Fig. 3.23.

3.6 Conclusions

The Papapetrou equations, which model a spinning test particle, exhibit chaotic solutions in Kerr spacetime for a wide range of parameters. In terms of the mass M of the central (Kerr) black hole, the largest Lyapunov exponents are of order $\lambda_{\max} = 0.01 M^{-1}$, which represents an exponential divergence of trajectories on a timescale of $\tau_\lambda = 1/\lambda = 100 M$. Furthermore, there are many chaotic orbits with exponents in the range 10^{-3} – $10^{-4} M^{-1}$. Despite the large number of chaotic orbits, we find that values of λ corresponding to unambiguous chaos occur exclusively when the spin parameter S is not small compared to unity. In particular, we find virtually no chaos for spin values below $S = 0.1$, and no evidence of any chaos for spins below $S = 10^{-4}$.

The strongest determinant of chaotic behavior, apart from the spin parameter, is the pericenter of the orbit in question. The most highly chaotic orbits are those that reach pericenters near the horizon of the black hole. This is due to the high spacetime curvature in these regions, which maximizes the size of the coupling of the spin to the Riemann curvature tensor [Eq. (3.1)]. When the Kerr parameter a is small, so that the Kerr metric differs only slightly from the Schwarzschild metric, the pericenters are much higher than in the extreme Kerr ($a = 1$) case. Chaos in the Papapetrou system is therefore weak when a is small. The prevalence of chaos is also dependent on orbital eccentricity. Near-circular ($e = 0.01$) orbits have many fewer regions of chaotic orbits than those with higher eccentricities ($e = 0.5$ or $e = 0.6$). This seems due primarily to the lower pericenters accessible to high-eccentricity orbits.

The dependence of the Lyapunov exponents on S is our most important result: in all cases

considered, physically realistic values of S (satisfying $S \ll 1$) are not chaotic. We have shown conclusively that the Papapetrou equations admit many solutions that are formally chaotic, but without exception such chaotic solutions occur only for relatively large values of S . Below the upper limit for physically realistic spins ($S \sim 10^{-4}$), we find no evidence of chaotic solutions. As a practical matter, this means that chaos will not manifest itself in the gravitational radiation from extreme mass-ratio binary inspirals.

Acknowledgements

I would like to thank Scott Hughes and Teviet Creighton for providing notes on parameterizing Kerr geodesics in terms of orbital parameters. I also thank Sterl Phinney for his careful reading of the manuscript and perceptive comments. This work was supported in part by NASA grant NAG5-10707.

Appendix: Spin vector formulation

We summarize here the formulation of the Papapetrou equations in terms of the spin 1-form S_μ (often referred to loosely as the “spin vector”), as mentioned in Sec. 3.2.1. In this chapter we consider a spinning particle of rest mass μ orbiting a central Kerr black hole of mass M , and it is convenient to measure all times and lengths in terms of M and all momenta in terms of μ . In these normalized units, the equations of motion in terms of the spin 1-form are

$$\begin{aligned} \frac{dx^\mu}{d\tau} &= v^\mu \\ \nabla_{\vec{v}} p_\mu &= -R_{\mu\nu}^*{}^{\alpha\beta} v^\nu p_\alpha S_\beta \\ \nabla_{\vec{v}} S_\mu &= -p_\mu \left(R^{*\alpha}{}_\beta{}^{\gamma\delta} S_\alpha v^\beta p_\gamma S_\delta \right) \end{aligned} \tag{3.58}$$

where

$$R^{*\alpha}{}_\beta{}^{\mu\nu} = \frac{1}{2} R^\alpha{}_{\beta\rho\sigma} \epsilon^{\rho\sigma\mu\nu}. \tag{3.59}$$

The supplementary condition Eq. (3.5) allows for an explicit solution for the 4-velocity v^μ in terms of p^μ :

$$v^\mu = N(p^\mu + w^\mu), \tag{3.60}$$

where

$$w^\mu = -{}^*R^{*\mu\alpha\beta\gamma} S_\alpha p_\beta S_\gamma \tag{3.61}$$

and

$${}^*R^{*\alpha\beta\mu\nu} = \frac{1}{2} R^{*\alpha\beta\rho\sigma} \epsilon_{\rho\sigma}{}^{\mu\nu}. \tag{3.62}$$

The normalization constant N is fixed by the constraint $v_\mu v^\mu = -1$.

The spin 1-form satisfies two orthogonality constraints:

$$p^\mu S_\mu = 0 \tag{3.63}$$

and

$$v^\mu S_\mu = 0. \tag{3.64}$$

These two constraints are equivalent as long as v^μ is given by Eq. (3.60): $0 = v^\mu S_\mu \propto p^\mu S_\mu + w^\mu S_\mu = p^\mu S_\mu$, since by definition of w^μ [Eq. (3.61)] the second term involves the contraction of a symmetric tensor with an antisymmetric tensor and therefore vanishes. We enforce Eq. (3.63) in our parameterization scheme, and we use Eq. (3.60) in the equations of motion, so Eq. (3.64) is then automatically satisfied.

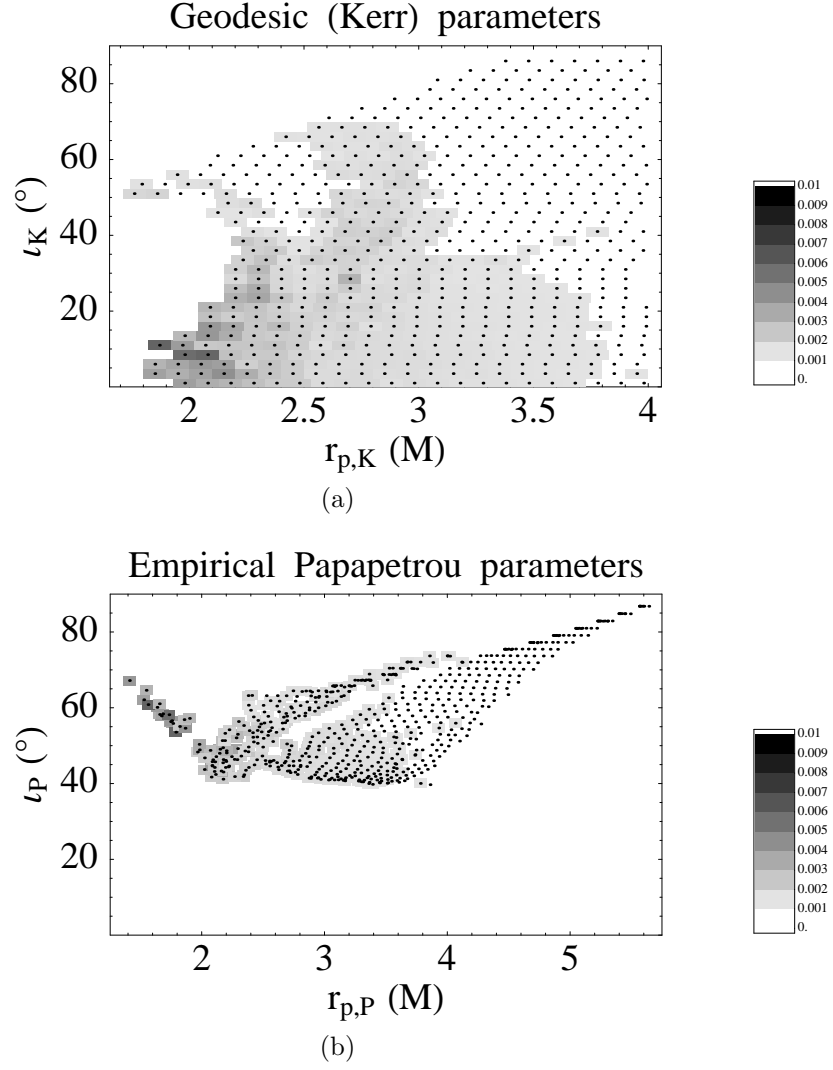


Figure 3.25: r_p - ι map for $S = 0.9$, $a = 1$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. Chaotic orbits are widespread.

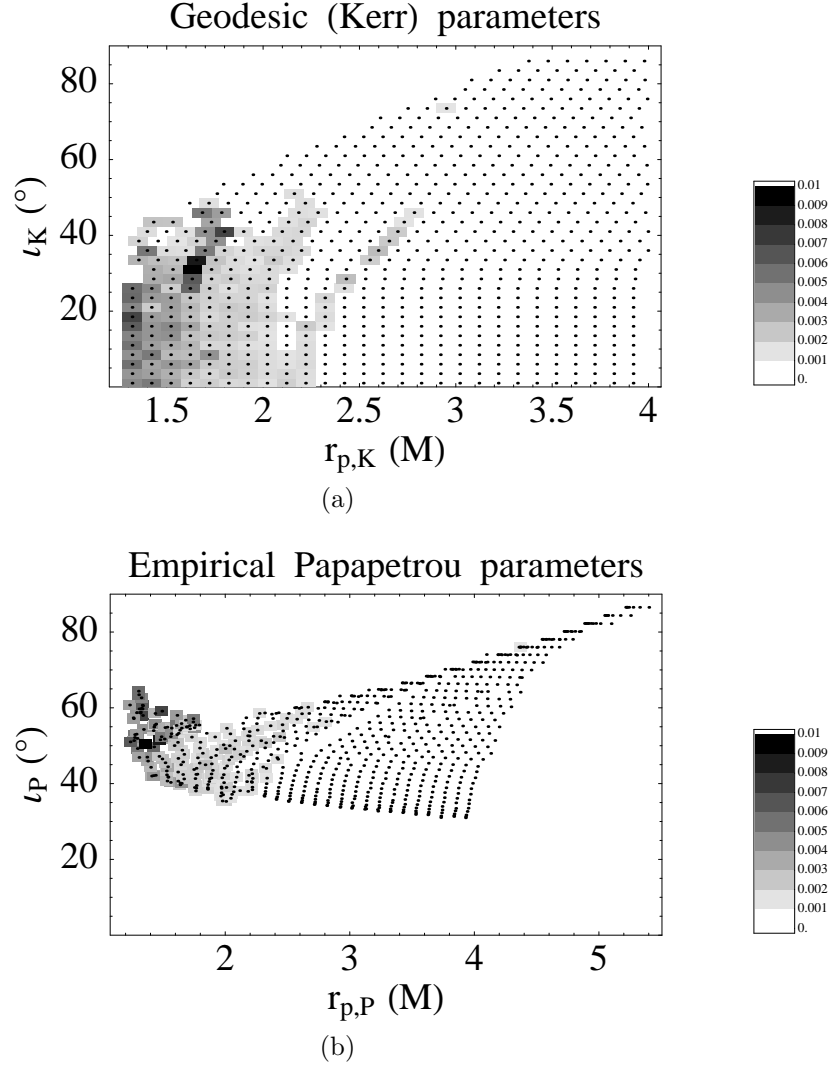


Figure 3.26: r_p - ι map for $S = 0.5$, $a = 1$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. Because of the extremely low pericenters accessible at this value of S (which are excluded when $S = 1$), the chaos for $S = 0.5$ is the strongest we find. The largest Lyapunov exponent is just over $\lambda = 0.01 M^{-1}$.

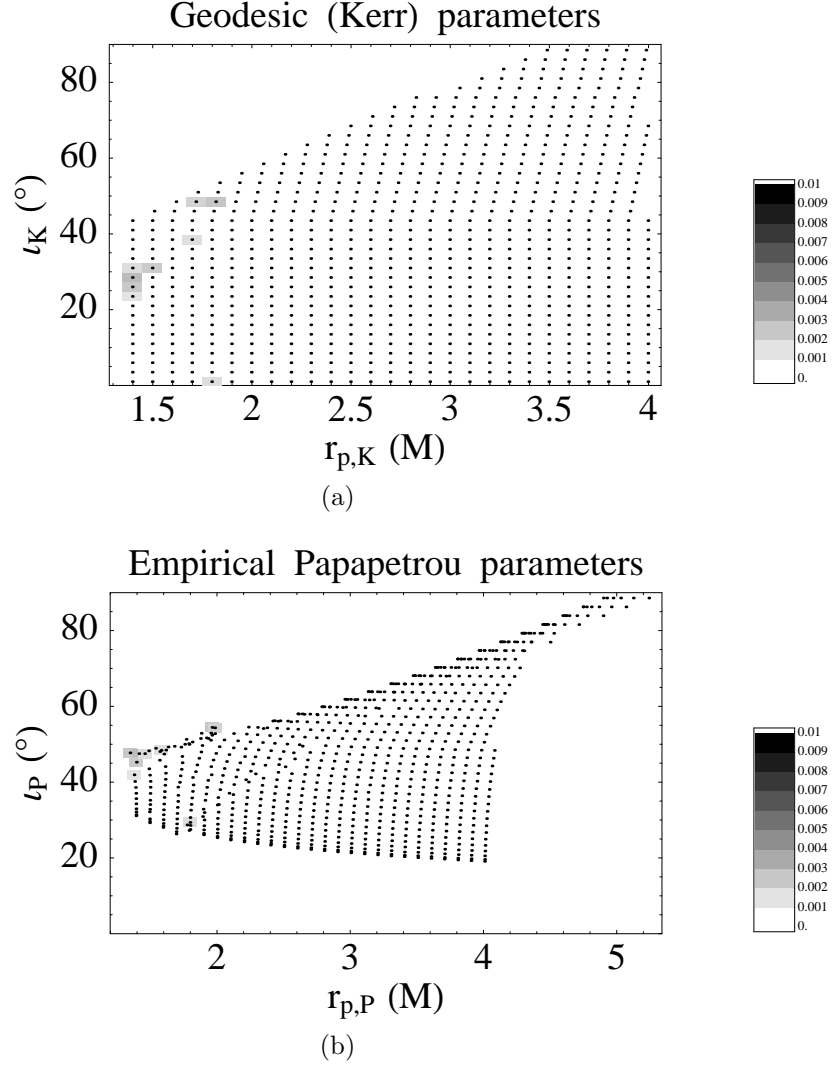


Figure 3.27: r_p - ι map for $S = 0.2$, $a = 1$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. Only a few orbits are chaotic.

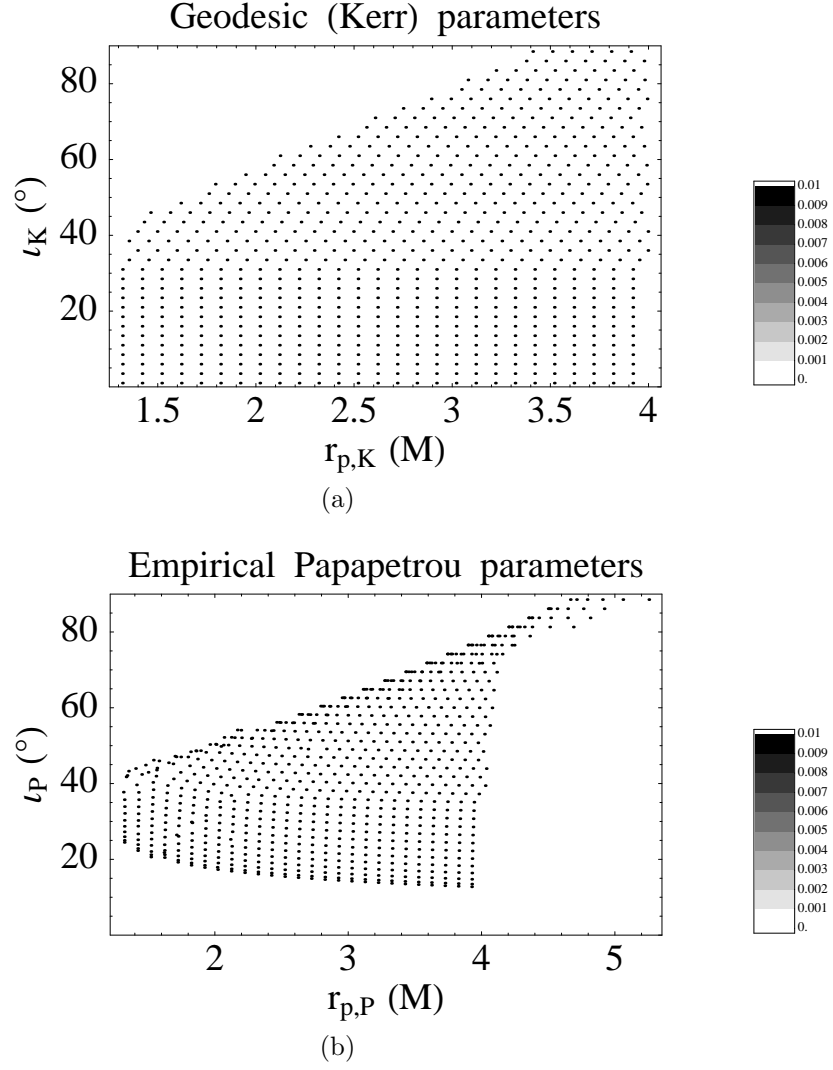


Figure 3.28: r_p - ι map for $S = 0.1$, $a = 1$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. All chaos is gone, although the parameter space is still somewhat compressed.

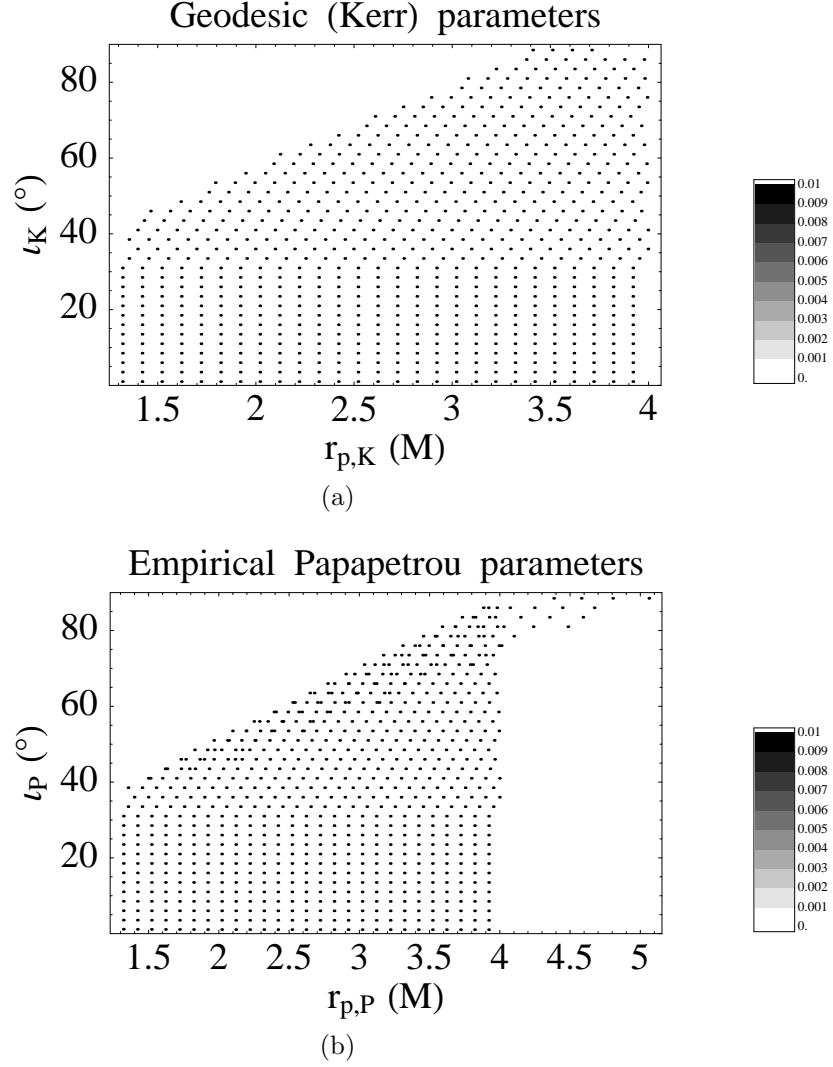


Figure 3.29: r_p - ι map for $S = 10^{-4}$, $a = 1$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. There are no chaotic orbits. The empirical parameter values in (b) are indistinguishable from the requested values except for initial conditions that specify values of ι corresponding to unstable orbits (as discussed in Sec. 3.3.2.3).

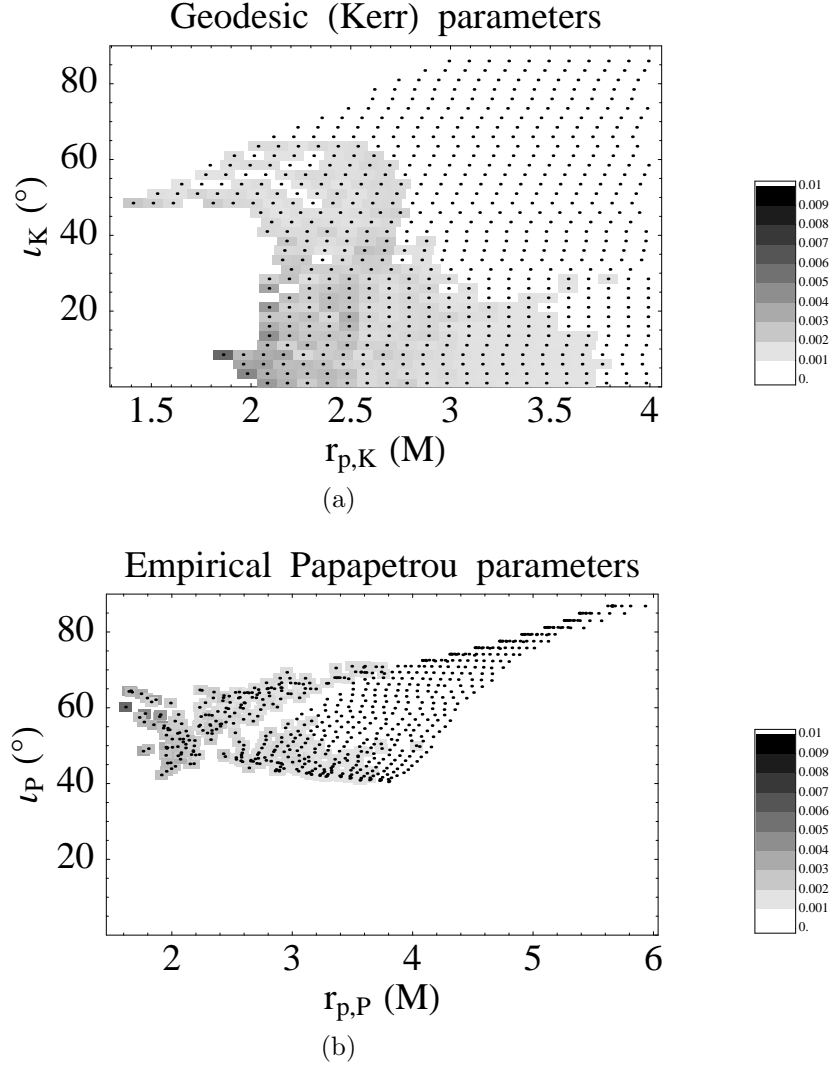


Figure 3.30: r_p - ι map for $S = 1$, $a = 1$, and $e = 0.6$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. Chaotic orbits are widespread. The largest Lyapunov exponent is $\lambda = 5.5 \times 10^{-3} M^{-1}$.

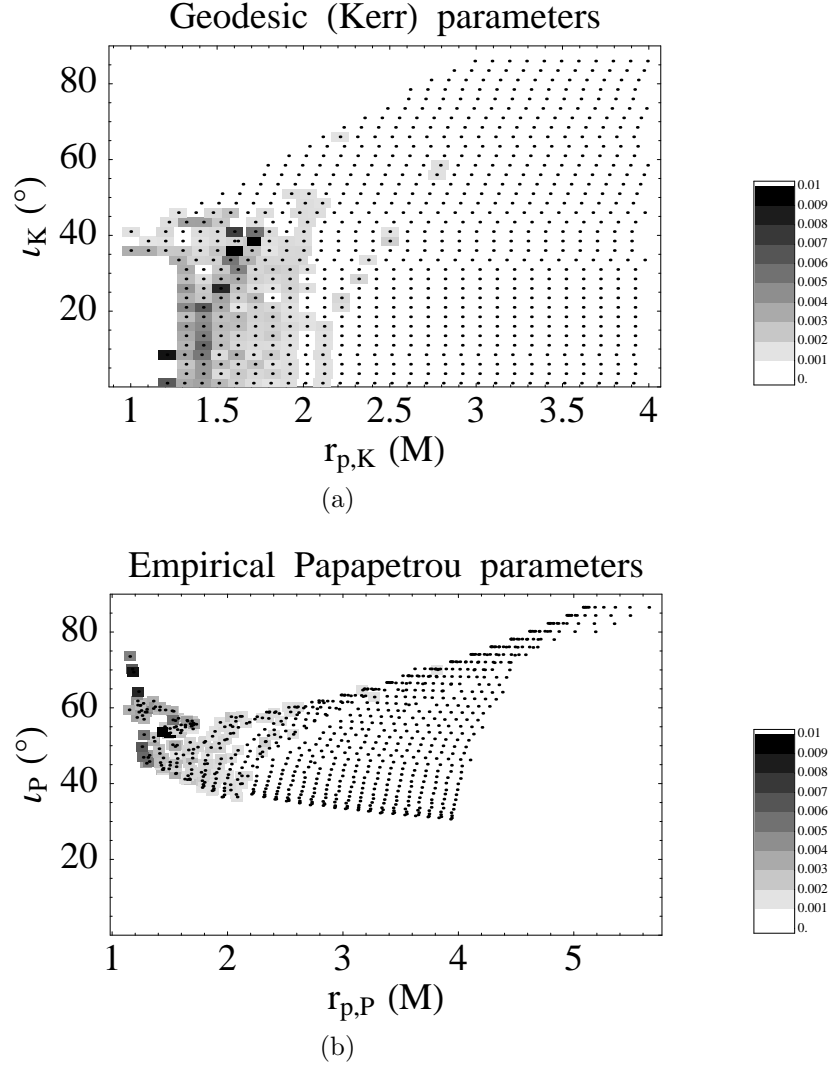


Figure 3.31: r_p - ι map for $S = 0.5$, $a = 1$, and $e = 0.6$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. As in the case of Fig. 3.26, the low accessible pericenters give rise to strong chaos. The largest Lyapunov exponent is $\lambda = 9.2 \times 10^{-3} M^{-1}$.

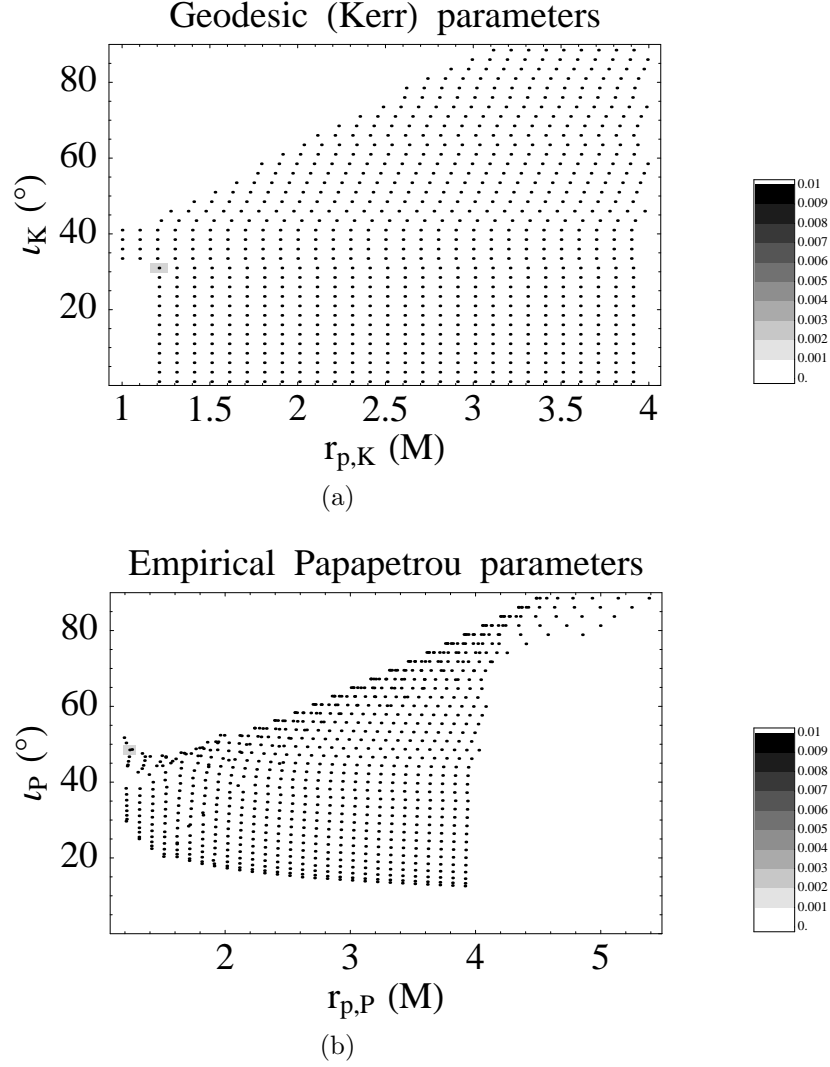


Figure 3.32: r_p - ι map for $S = 0.1$, $a = 1$, and $e = 0.6$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. There is only one chaotic initial condition (which is in fact the only $S = 0.1$ chaos we find), but the chaos is real, as discussed in Sec. 3.5.2 and illustrated in Fig. 3.12. The Lyapunov exponent is $\lambda = 1.0 \times 10^{-3} M^{-1}$.

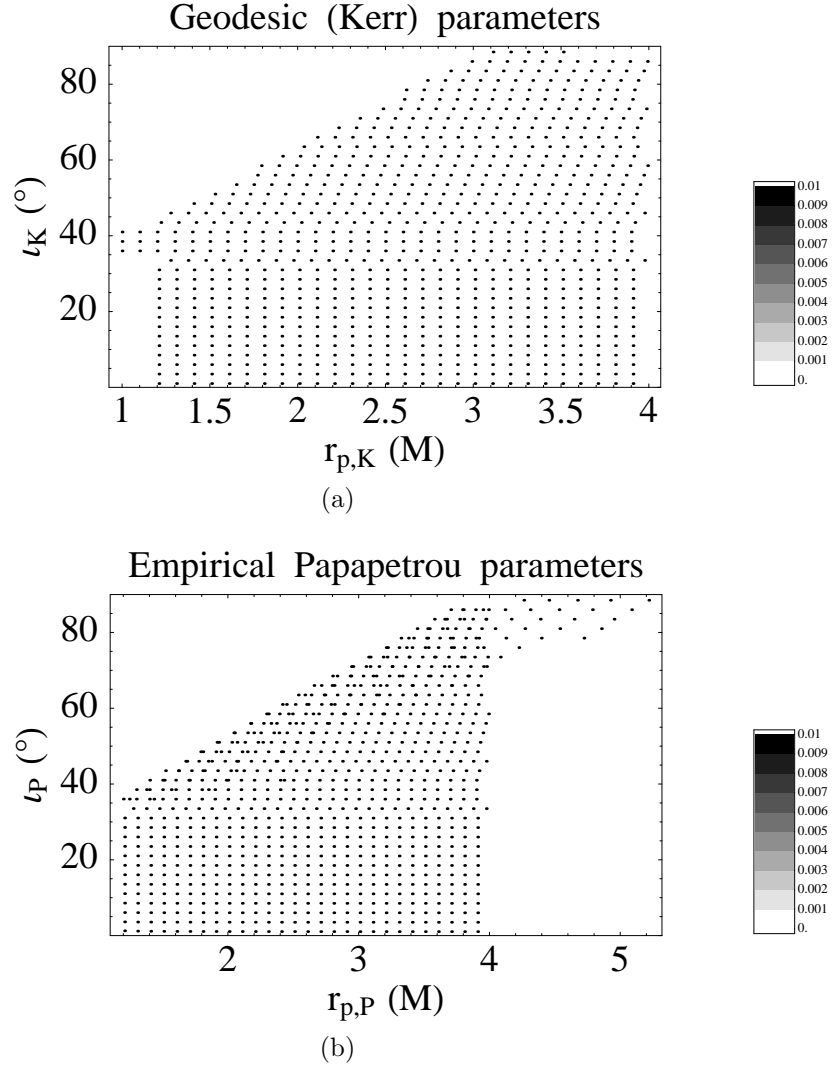


Figure 3.33: r_p - l map for $S = 10^{-4}$, $a = 1$, and $e = 0.6$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. All chaos has disappeared. As in Fig. 3.29, the empirical parameter values in (b) are indistinguishable from the requested values except for initial conditions that specify unstable orbits (Sec. 3.3.2.3).

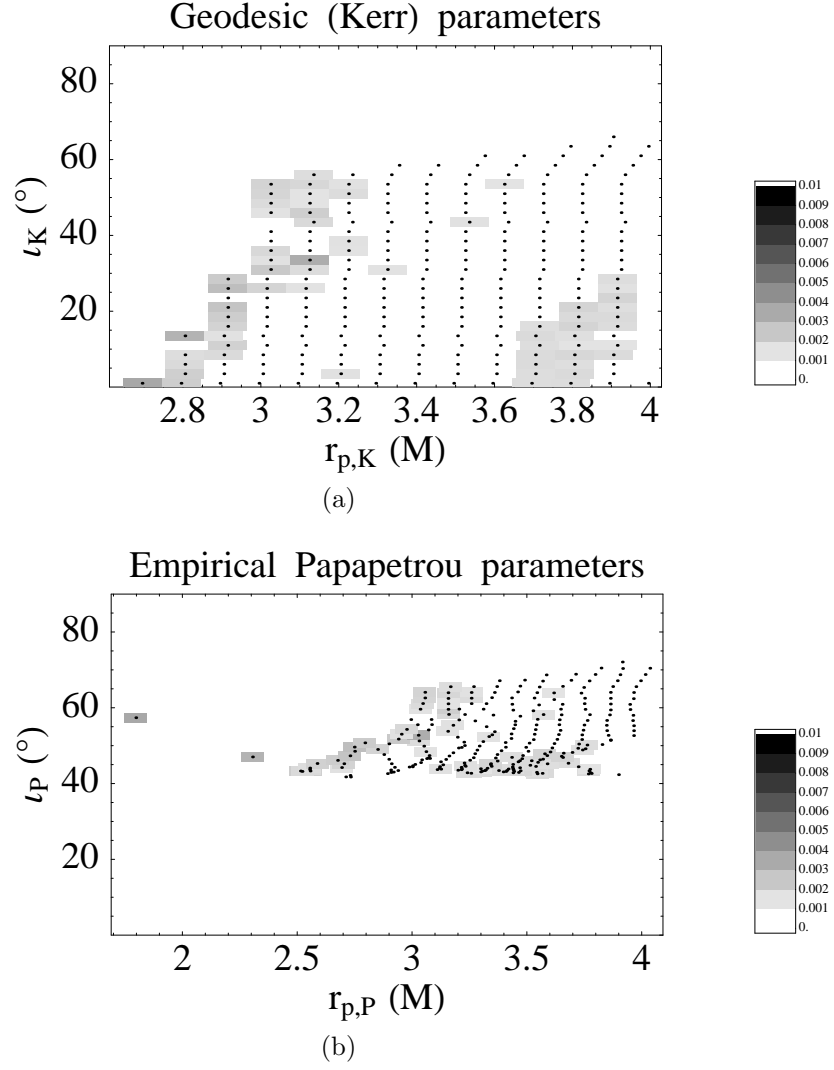


Figure 3.34: r_p - ι map for $S = 1$, $a = 1$, and $e = 0.01$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. There is some relatively weak chaos. The largest Lyapunov exponent is $\lambda = 2.7 \times 10^{-3} M^{-1}$.

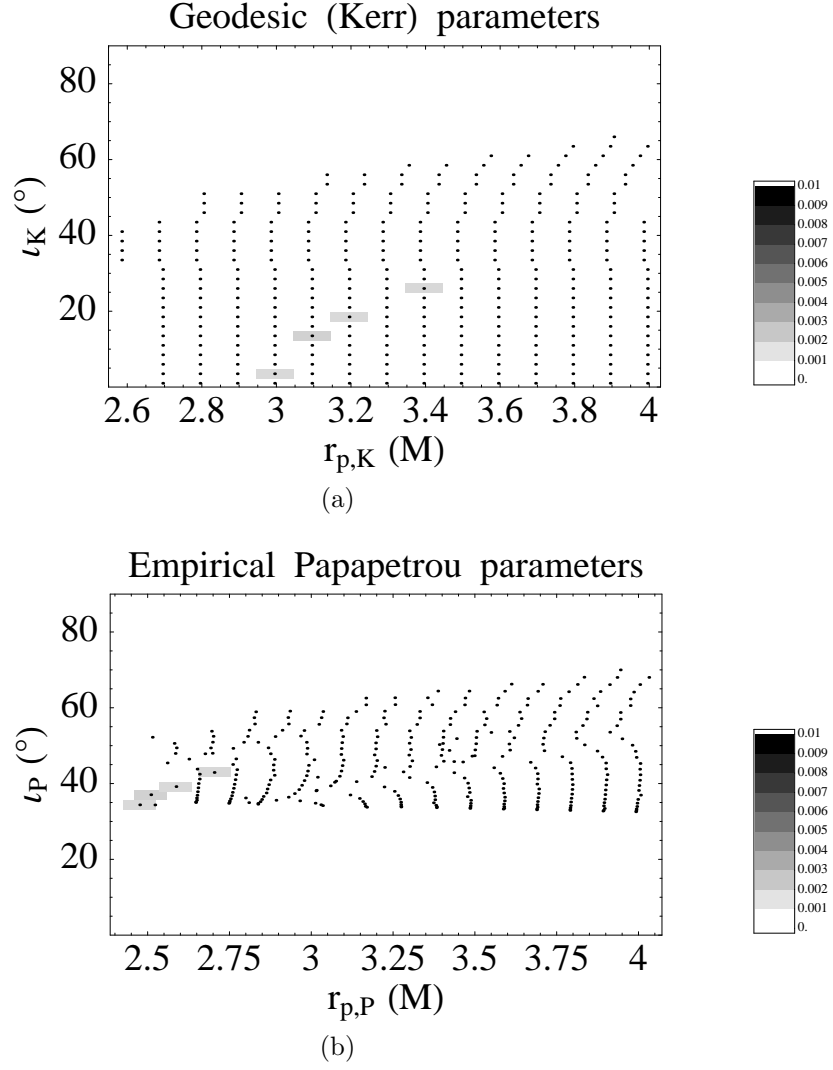


Figure 3.35: r_p - ι map for $S = 0.5$, $a = 1$, and $e = 0.01$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. There are a few regions of weak chaos.

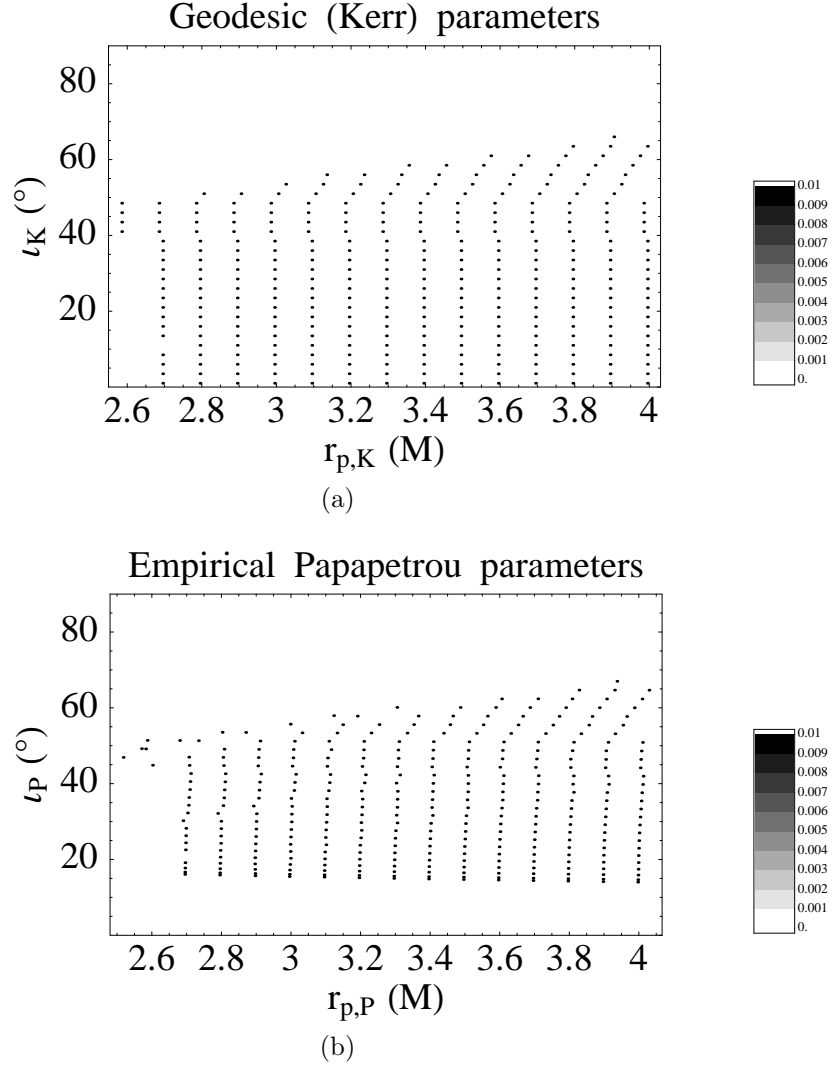


Figure 3.36: r_p - λ map for $S = 0.1$, $a = 1$, and $e = 0.01$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. There are apparently no chaotic initial conditions.

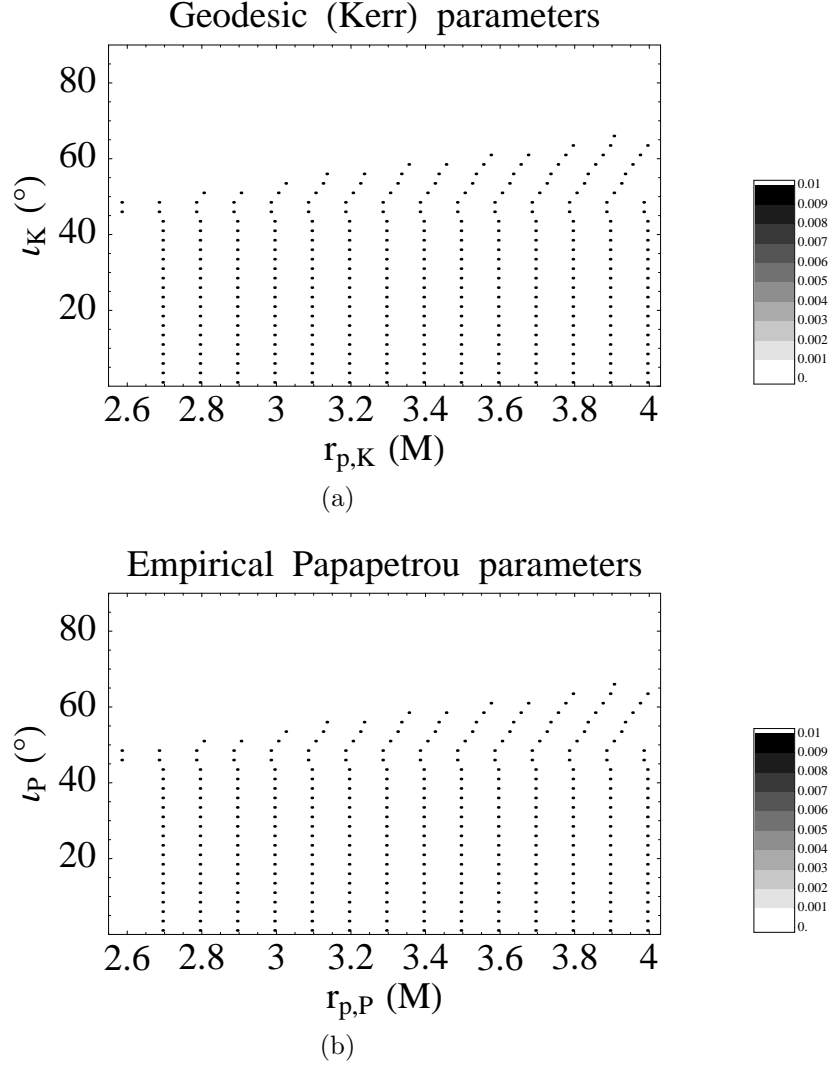


Figure 3.37: r_p - ι map for $S = 10^{-4}$, $a = 1$, and $e = 0.01$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. The chaos had disappeared.

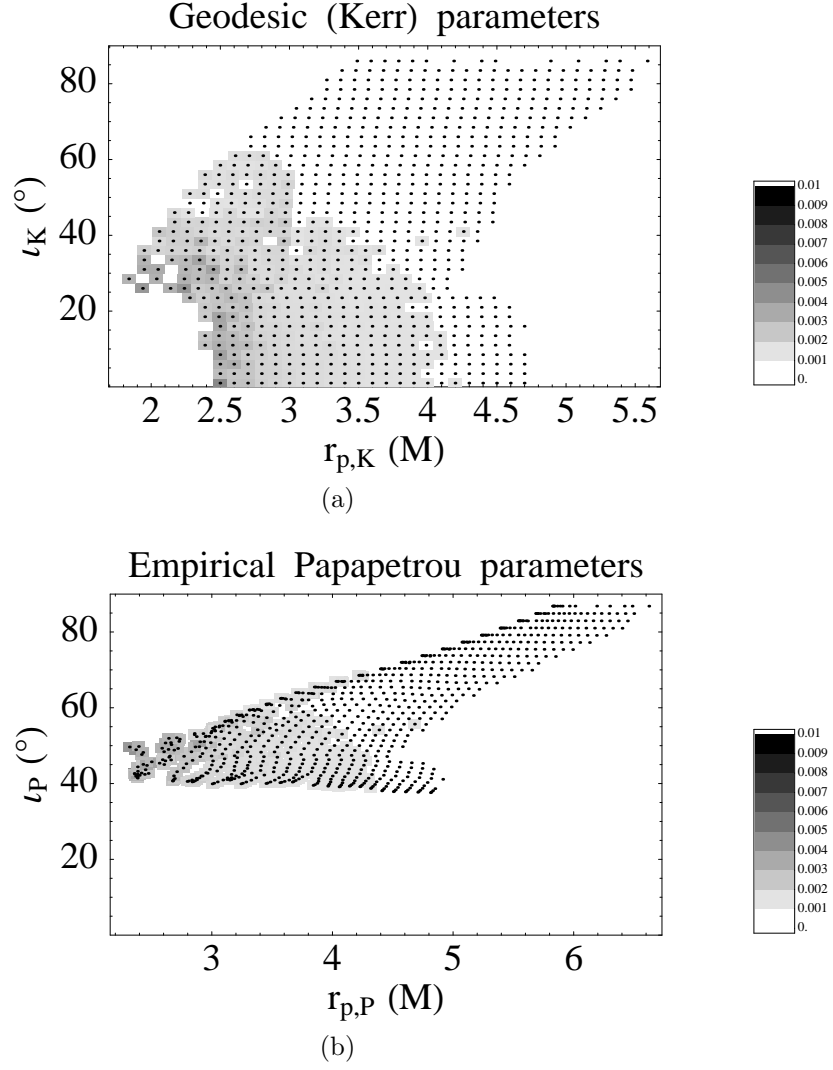


Figure 3.38: r_p - ι map for $S = 1$, $a = 0.9$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. Chaotic orbits are widespread.

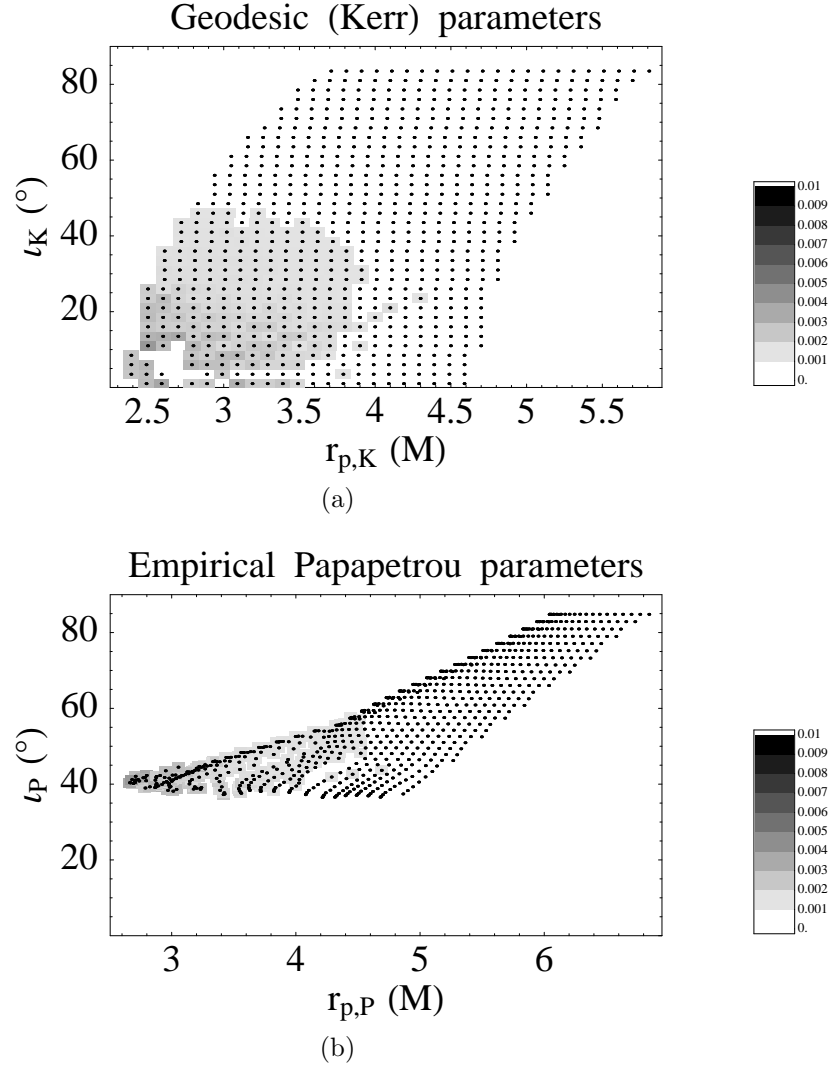


Figure 3.39: r_p - ι map for $S = 1$, $a = 0.7$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. There is still substantial chaos.

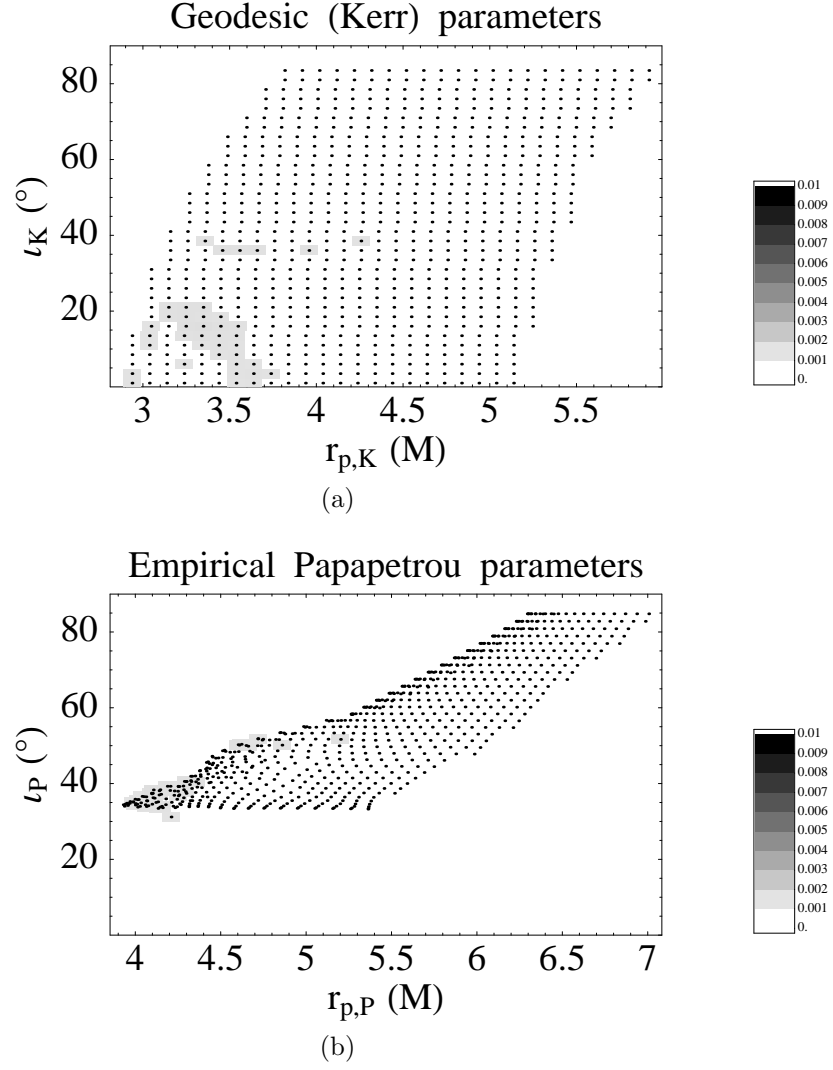


Figure 3.40: r_p - ι map for $S = 1$, $a = 0.5$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. The chaos is largely confined to low pericenter orbits.

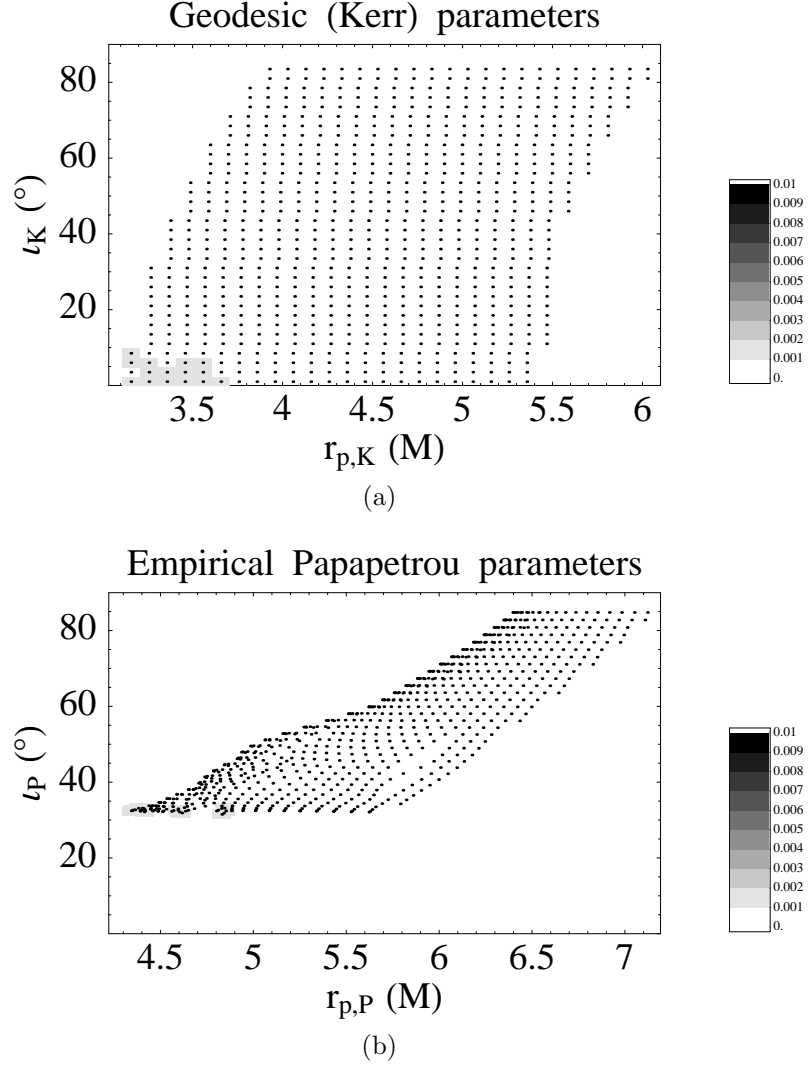


Figure 3.41: r_p - ι map for $S = 1$, $a = 0.4$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. Only a handful of initial conditions are chaotic.

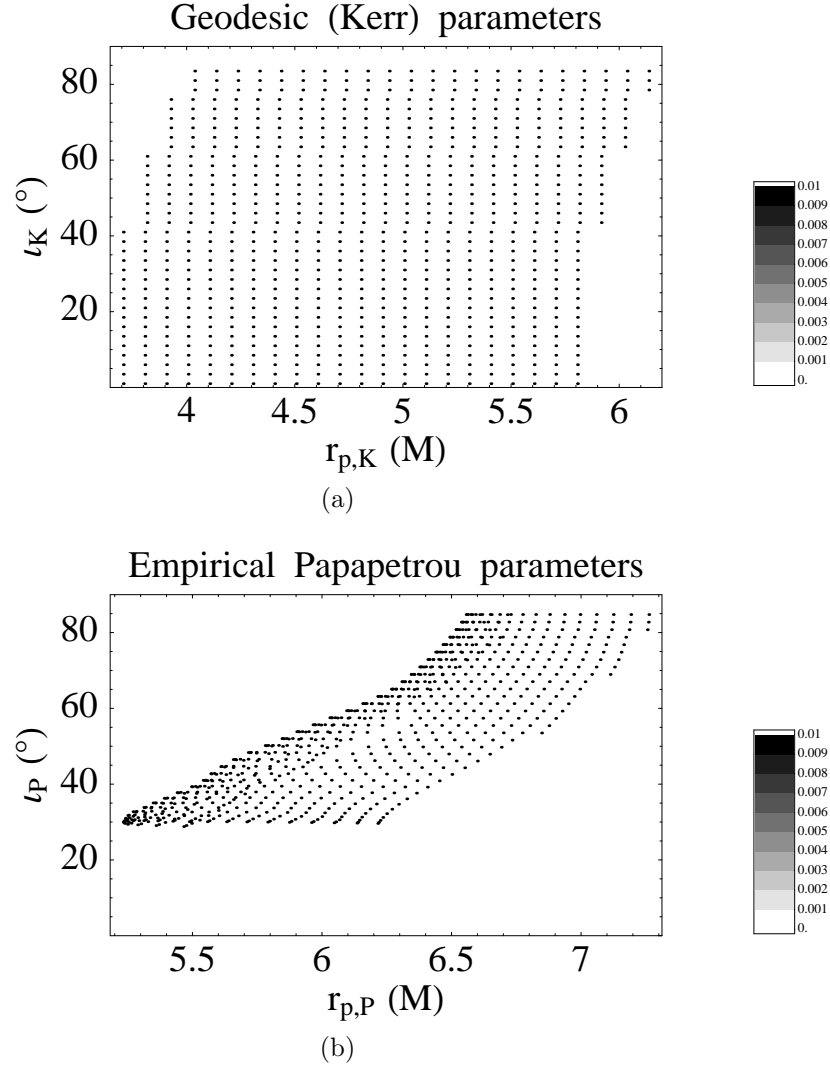


Figure 3.42: r_p - ι map for $S = 1$, $a = 0.2$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. No initial conditions are chaotic.

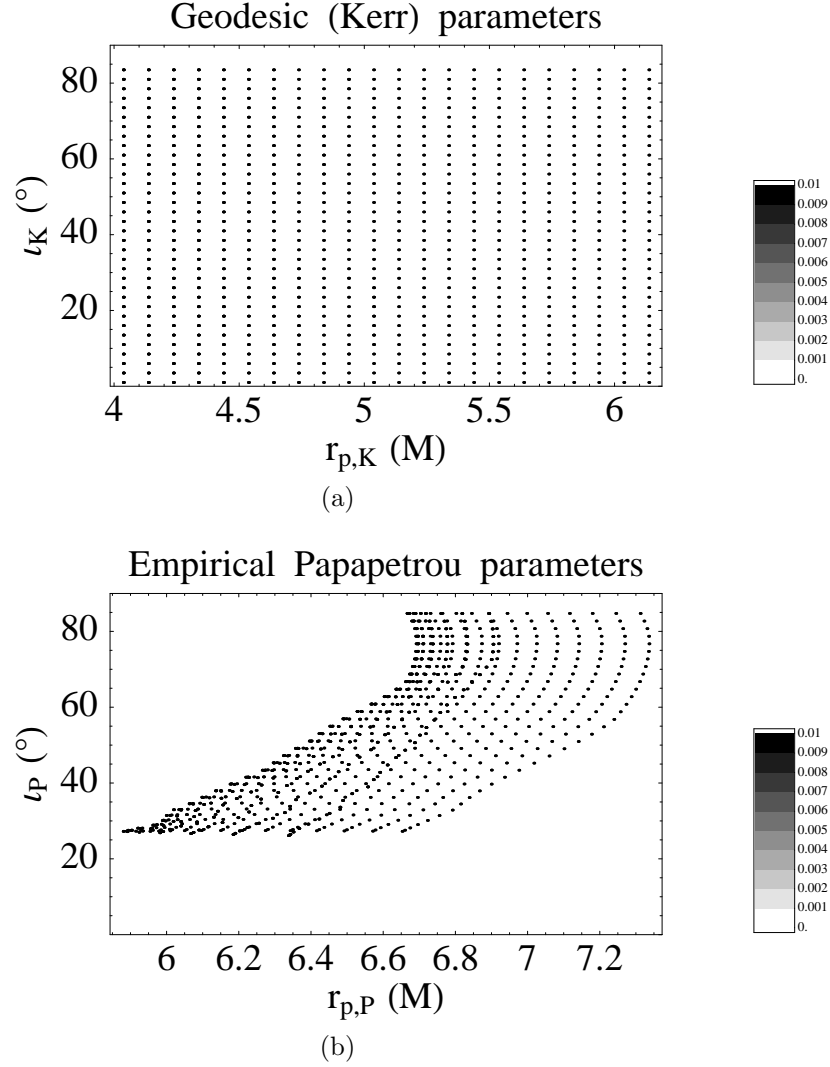


Figure 3.43: r_p - ι map for $S = 1$, $a = 0$, and $e = 0.5$. (a) Requested parameters; (b) empirical parameters. The shading is scaled to the same maximum Lyapunov exponent as in Fig. 3.9. No initial conditions are chaotic. Note that every column of (a) is identical. This is a result of the spherical symmetry of the $a = 0$ (Schwarzschild) metric: all inclination angles are equivalent. As seen in (b), this symmetry is broken by the spin of the test particle.

Bibliography

- [1] S. Suzuki and K. Maeda, Phys. Rev. D **55**, 4848 (1997).
- [2] J. Levin, Phys. Rev. Lett. **84**, 3515 (2000).
- [3] J. Levin, Phys. Rev. D **67**, 044013 (2003).
- [4] N. J. Cornish, Phys. Rev. D **64**, 084011 (2001).
- [5] N. J. Cornish and J. Levin, Phys. Rev. Lett. **89**, 179001 (2002).
- [6] N. J. Cornish and J. Levin, [gr-qc/0207016](#).
- [7] T. Tanaka, Y. Mino, M. Sasaki, and M. Shibata, Phys. Rev. D **54**, 3762 (1996).
- [8] M. D. Hartl, Phys. Rev. D **67**, 024005 (2003), [gr-qc/0210042](#).
- [9] O. Semerák, Mon. Not. R. Astron. Soc. **308**, 863 (1999).
- [10] S. Suzuki and K. Maeda, Phys. Rev. D **58**, 023005 (2000), [gr-qc/9712095](#).
- [11] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
- [12] A. Papapetrou, Proc. R. Soc. London **A209**, 248 (1951).
- [13] W. G. Dixon, Proc. R. Soc. London **A314**, 499 (1970).
- [14] S. Hughes, http://www.tapir.caltech.edu/listwg1/EMRI/hughes_geodesic.html.
- [15] F. D. Ryan, Phys. Rev. D **53**, 3064 (1996).
- [16] R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago, 1984).
- [17] V. Karas and D. Vokrouhlický, Gen. Relativ. and Grav. **24**, 729 (1992).
- [18] S. A. Hughes, Phys. Rev. D **64**, 064004 (2001).
- [19] <http://lisa.jpl.nasa.gov>.

Chapter 4

Dynamics of binary black holes using the post-Newtonian approximation

This chapter represents ongoing research based in part on suggestions by Alessandra Buonanno, and is expected to continue (in expanded form) in collaboration with her.

Abstract

We investigate the dynamics of binary black holes using an effective one-body, Hamiltonian formulation of the post-Newtonian (PN) equations of motion. The Hamiltonian includes four distinct spin effects: spin-orbit coupling, spin-spin coupling, and two mass monopole/spin-induced quadrupole interaction terms. We investigate the qualitative effects of these terms on the orbits, and we also compare post-Newtonian orbits with analogous solutions of the Papapetrou equations, which model a spinning particle in the extreme mass-ratio limit. In the special case of quasi-circular orbits, we search for the presence of chaos (using the method of Lyapunov exponents) for a large variety of initial conditions. We find that the addition of the spin-spin terms leads to strongly chaotic behavior, but the chaos is rare, occurring only in a tiny fraction of cases. Moreover, chaos occurs only for orbits with low radii, where the approximations required for the physical validity of the post-Newtonian equations are not well satisfied.

4.1 Introduction

Relativistic binary systems made of compact objects, such as neutron stars or black holes, are among the most promising (or, at least, are among the best-understood) candidates for the production of gravitational waves detectable by both ground- and space-based gravitational wave observatories. The difficulty of detecting the signals from such systems has led to a theoretical effort to under-

stand the gravitational waveforms likely to be emitted by such systems, which in turn has led to a consideration of their dynamical behavior. In particular, several authors [1–5] have investigated the presence of chaos in the dynamics of compact binaries, motivated in part by the effect chaos could have on the calculation of theoretical templates for use in matched filters [6–8]. The extreme sensitivity on initial conditions that characterizes chaotic systems would lead to significant difficulties in the implementation of such filters, since the number of templates required for a good match would grow exponentially with increasing detection sensitivity [9].

In the Papapetrou equations, which model the dynamics of a spinning test particle, chaos was found for the case of a spinning particle orbiting a nonrotating (Schwarzschild) black hole [1]. We extended this result to the case of a rotating (Kerr) black hole in Chapters 2 and 3, finding widespread chaotic solutions. As demonstrated in Chapter 2, however, the values of the total spin for the test particle leading to chaotic solutions are not realizable in physical systems that satisfy the approximations made in deriving the Papapetrou equations. Furthermore, in Chapter 3 we showed that chaos, while widespread for these unrealistic spin values, disappears in all cases for physically realistic spins. In short, there is strong evidence that extreme mass-ratio systems, which are most relevant for proposed space-based gravitational wave detectors, are not chaotic for any parameter values of physical interest.

The case of comparable-mass binaries has been investigated by several authors [2–4] using the post-Newtonian (PN) equations of motion (Sec. 4.2). There was initially some doubt regarding the results presented in [2], which found chaos in the PN equations for spinning bodies, since the timescale of the chaos was not reported: it was not clear whether the chaos discovered in the equations—as calculated in the conservative limit neglecting gravitational radiation reaction—would have time to manifest itself in the inspiral timescale t_{insp} . The work in [3] cast doubt on the presence of chaos in these systems, finding that the Lyapunov characteristic exponents for the PN equations, which measure the divergence rate of nearby trajectories, are zero in all cases tested. However, [4] found some initial conditions that do have positive Lyapunov exponents, indicating the presence of chaos, with characteristic times shorter than t_{insp} , thus raising the possibility that theoretical templates calculated for systems with spinning compact objects are affected by chaos.

In the present study, we examine and extend these results by investigating the dynamics of spinning binary black holes using a Hamiltonian formulation of the post-Newtonian equations of motion. In order to make chaos formally possible, we exclude gravitational radiation reaction—since tests for chaos technically require an infinite-time limit, the finite inspiral times due to radiation reaction would eliminate the possibility of chaos. On the other hand, we do include post-Newtonian terms involving the spin of the two bodies: the addition of spin is essential to create the possibility of chaos, since without spin the conserved quantities constrain the motion to be at worst quasi-periodic. As discussed in Sec. 4.2.1, we use four separate spin terms in the equations of motion

to model accurately their effect of the dynamics. We focus on black holes, to the exclusion of other compact astrophysical objects, because two of these spin terms (which involve spin-induced quadrupole effects) are known exactly only for black holes, and yet their magnitudes are comparable to the spin-spin coupling and hence cannot be ignored. (This is an extension of previous work, as other authors have not considered these quadrupole terms when investigating chaos.)

In Sec. 4.4, we examine the PN dynamics in the extreme mass-ratio limit, in order to make contact with Chapters 2 and 3. We then investigate chaos for comparable-mass binary black holes (Sec. 4.5). Since binary black hole inspirals tend to *circularize* under gravitational radiation reaction, we focus on the important special case of *quasi-circular* orbits.¹ As in previous work, we favor Lyapunov exponents (Sec. 4.5.1) to quantify the presence (or absence) of chaos.

We work almost exclusively in geometric units ($G = c = 1$). Euclidean vectors, such as those that appear in the post-Newtonian equations of motion, are set in boldface, and we use vector arrows to denote relativistic 4-vectors. The symbol \log refers in all cases to the natural logarithm.

4.2 The post-Newtonian equations of motion

The post-Newtonian (PN) equations for the two-body problem are an approximation to full general relativity, essentially involving a series expansion in v/c . The PN equations are therefore valid in the limit $v \ll c$; since (in geometric units) the velocity satisfies $v^2 \sim M/r$, where M is the total mass of the binary, the condition $v \ll 1$ requires that the radius satisfy $r \ll M$. As in the case of the classical two-body problem, in the post-Newtonian case it is possible to reduce the motion of a relativistic binary to a one-body problem. A typical orbit using the one-body formulation is shown in Fig. 4.1. In this chapter, we use an effective one-body, Hamiltonian approach, as developed in [10], which provides a particularly clean derivation of the equations of motion. In addition, since detecting chaos involves determining the separation of nearby *phase-space* trajectories, is it convenient to work directly in terms of spatial coordinates and their corresponding conjugate momenta—a criterion automatically satisfied by the Hamiltonian formulation.

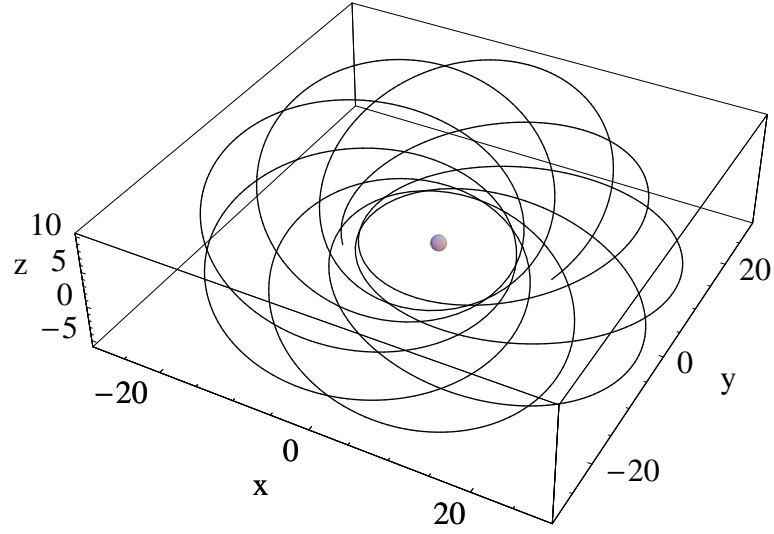
4.2.1 The Hamiltonian formulation

We can represent the PN Hamiltonian schematically as follows [10–12]:

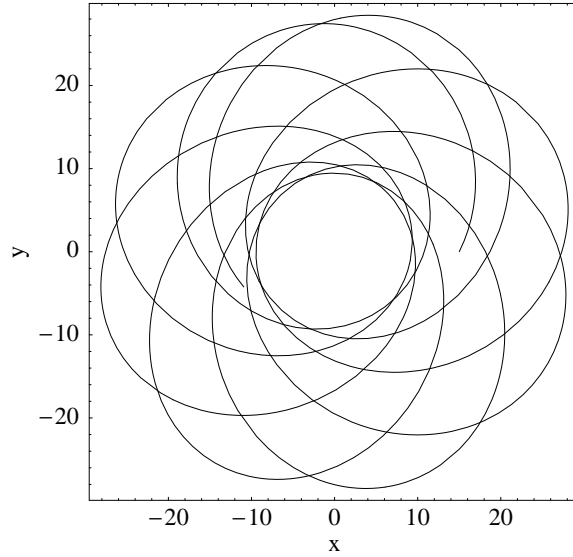
$$H = H_N + H_{\text{PN}} + H_{\text{SO}} + H_{\text{SS}}. \quad (4.1)$$

We include the following terms: Newtonian, standard post-Newtonian (through 2PN, i.e., v^4/c^4), spin-orbit coupling, and spin-spin coupling. (We omit the radiation reaction, as discussed in the

¹A thorough investigation of eccentric orbits is more involved due to the large number of possible orbits, and is left for future work.



(a)



(b)

Figure 4.1: The orbit of two maximally spinning $10 M_{\odot}$ black holes, using a Hamiltonian, effective one-body formulation of the post-Newtonian equations of motion. (a) The orbit embedded in Euclidean space; (b) the projection onto the x - y plane. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$, indicating the collapse radius where the relative separation of the two bodies is the sum of their horizon radii. (We have $r_{H,1} = m_1$, and $r_{H,2} = m_2$ for maximally spinning black holes, so the relative separation of collapse is $r_H = m_1 + m_2 = M$.) Note that, in contrast to Newtonian orbits, the orbit is not closed, and the orbital plane precesses around the center of mass. The initial conditions producing this orbit are $\mathbf{X} = (15, 0, 0)$, $\mathbf{P} = (0.137306, 0.262401, 0.0929462)$, $\mathbf{S}_1 = (0, 0, 1)$, and $\mathbf{S}_2 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}})$, with all the PN terms from Sec. 4.2.1 turned on.

introduction.) Throughout this treatment, we use \mathbf{X} for the (relative) position, \mathbf{P} for the conjugate (relative) momenta, and $(\mathbf{S}_1, \mathbf{S}_2)$ for the spins of the two particles.

For two bodies with masses m_1 and m_2 , the effective one-body approach introduces the total mass $M = m_1 + m_2$ and the reduced mass $\mu = m_1 m_2 / M$ (precisely as in the solution of the Newtonian two-body problem).² Using these mass variables, we can express the first term (the standard Newtonian energy) as follows:

$$H_N = \frac{P^2}{2\mu} - \frac{\mu M}{r}, \quad (4.2)$$

where $r = \|\mathbf{X}\|$. The post-Newtonian terms used in this chapter are 1PN and 2PN, given by

$$H_{\text{PN}} = \mu(\hat{H}_{1\text{PN}} + \hat{H}_{2\text{PN}}), \quad (4.3)$$

where

$$\hat{H}_{1\text{PN}} = \frac{1}{8}(3\eta - 1)(\mathbf{p}^2)^2 - \frac{1}{2}[(3 + \eta)\mathbf{p}^2 + \eta(\mathbf{n} \cdot \mathbf{p})^2]\frac{1}{q} + \frac{1}{2q^2} \quad (4.4)$$

and

$$\begin{aligned} \hat{H}_{2\text{PN}} = & \frac{1}{16}(1 - 5\eta + 5\eta^2)(\mathbf{p}^2)^3 + \frac{1}{8}[(5 - 20\eta - 3\eta^2)(\mathbf{p}^2)^2 - 2\eta^2(\mathbf{n} \cdot \mathbf{p})^2\mathbf{p}^2 \\ & - 3\eta^4(\mathbf{n} \cdot \mathbf{p})^4]\frac{1}{q} + \frac{1}{2}[(5 + 8\eta)(\mathbf{p}^2) + 3\eta(\mathbf{n} \cdot \mathbf{p})^2]\frac{1}{q^2} - \frac{1}{4}(1 + 3\eta)\frac{1}{q^3}, \end{aligned} \quad (4.5)$$

and we use the unit vector $\mathbf{n} = \mathbf{X}/r$, the auxiliary variable $\eta = \mu/M$, and the reduced canonical variables $\mathbf{p} = \mathbf{P}/\mu$ and $\mathbf{q} = \mathbf{X}/M$. In most of this chapter, we will measure momenta in terms of μ and distances in terms of M , so we will typically not distinguish between the canonical and reduced canonical variables.

The next term in Eq. (4.1) is the spin-orbit coupling (corresponding to Lense-Thirring precession in the extreme mass-ratio limit $m_1 \gg m_2$):

$$H_{\text{SO}} = \frac{\mathbf{L} \cdot \mathbf{S}_{\text{eff}}}{r^3}, \quad (4.6)$$

where

$$\mathbf{S}_{\text{eff}} = \left(2 + \frac{3}{2} \frac{m_2}{m_1}\right) \mathbf{S}_1 + \left(2 + \frac{3}{2} \frac{m_1}{m_2}\right) \mathbf{S}_2. \quad (4.7)$$

We have three separate terms quadratic in the spin, so that the last term in Eq. (4.1) is

$$H_{\text{SS}} = H_{\text{S}_1\text{S}_2} + H_{\text{S}_1\text{S}_1} + H_{\text{S}_2\text{S}_2}. \quad (4.8)$$

²For clarity in what follows, we adopt the arbitrary convention that $m_1 \geq m_2$.

The first term, the spin-spin coupling, is

$$H_{S_1 S_2} = \frac{1}{r^3} [3(\mathbf{S}_1 \cdot \mathbf{n})(\mathbf{S}_2 \cdot \mathbf{n}) - \mathbf{S}_1 \cdot \mathbf{S}_2], \quad (4.9)$$

which is valid for all bodies (e.g., neutron stars or white dwarfs, as well as black holes). The next two terms we include are monopole-quadrupole interaction terms, and their form is valid only for black holes.³ They are

$$H_{S_1 S_1} = \frac{1}{2r^3} [3(\mathbf{S}_1 \cdot \mathbf{n})(\mathbf{S}_1 \cdot \mathbf{n}) - \mathbf{S}_1 \cdot \mathbf{S}_1] \frac{m_2}{m_1} \quad (4.10)$$

and

$$H_{S_2 S_2} = \frac{1}{2r^3} [3(\mathbf{S}_2 \cdot \mathbf{n})(\mathbf{S}_2 \cdot \mathbf{n}) - \mathbf{S}_2 \cdot \mathbf{S}_2] \frac{m_1}{m_2}. \quad (4.11)$$

With the full Hamiltonian in hand, we can now derive the equations of motion using Poisson brackets. As in classical Hamiltonian mechanics, the time-evolution of a dynamical quantity $f(\mathbf{X}, \mathbf{P}, \mathbf{S}_1, \mathbf{S}_2)$ is simply the Poisson bracket of the quantity with the Hamiltonian:

$$\frac{df}{dt} = \{f, H\}. \quad (4.12)$$

The Hamiltonian equations of motion for the (relative) position and (relative) momentum are then the familiar canonical equations:

$$\frac{d\mathbf{X}}{dt} = +\frac{\partial H}{\partial \mathbf{P}}, \quad \frac{d\mathbf{P}}{dt} = -\frac{\partial H}{\partial \mathbf{X}}. \quad (4.13)$$

To derive the spin equations of motion, we use the canonical angular momentum Poisson bracket

$$\{S^i, S^j\} = \epsilon^{ijk} S^k, \quad (4.14)$$

which yields

$$\frac{d\mathbf{S}_1}{dt} = \frac{\partial H}{\partial \mathbf{S}_1} \times \mathbf{S}_1 \equiv \boldsymbol{\Omega}_1 \times \mathbf{S}_1 \quad (4.15)$$

and

$$\frac{d\mathbf{S}_2}{dt} = \frac{\partial H}{\partial \mathbf{S}_2} \times \mathbf{S}_2 \equiv \boldsymbol{\Omega}_2 \times \mathbf{S}_2. \quad (4.16)$$

Eqs. (4.13) and (4.15)–(4.16), with the Hamiltonian given by Eq. (4.1), are the equations of motion used throughout this chapter.

³We leave to future work the generalization to neutron stars and other compact bodies whose quadrupole moments depend on their internal structure and equations of state.

4.2.2 Conserved quantities

There are many conserved quantities in the post-Newtonian equations. These constants of the motion constrain the dynamical behavior of the system and provide valuable checks when testing a numerical implementation of the equations. Here we discuss all the quantities known to be conserved, and at which orders they are conserved.

4.2.2.1 Quantities conserved at all orders

The following quantities are conserved at all orders:

- Total energy H : $\dot{H} = \{H, H\} = 0$ by the antisymmetry of Poisson brackets.
- Total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}_1 + \mathbf{S}_2$: see [10].
- The spin magnitudes S_1 and S_2 : this is evident from Eqs. (4.15) and (4.16), since the cross product is perpendicular to the spin and hence can change only its direction.

4.2.2.2 Quantities conserved only through spin-orbit coupling

If we neglect the terms quadratic in the spin (i.e., we include only terms through spin-orbit coupling), the following additional quantities are conserved:

- $L^2 = \mathbf{L} \cdot \mathbf{L}$: at this order, $\dot{\mathbf{L}} = \mathbf{S}_{\text{eff}} \times \mathbf{L}/r^3$, which changes the direction of \mathbf{L} but not its magnitude.
- $\mathbf{L} \cdot \mathbf{S}_{\text{eff}}$: see [10].

In our numerical implementation (Sec. 4.2.3), we verify that the correct quantities are conserved at the proper orders, to within the accuracy of the numerical integrations.

4.2.3 Numerical implementation

Our primary implementation of the post-Newtonian equations of motion is a collection *Mathematica* packages, relying on the native `NDSolve` function to effect numerical integrations. A primary virtue of *Mathematica* is the ability to type in the Hamiltonian directly, with the equations of motion calculated automatically using *Mathematica*'s symbolic differentiation capabilities. From a debug/development perspective, the typesetting capabilities of *Mathematica* make it particularly easy to inspect the equations visually for consistency with printed representations. We implement the equations in a standard way to eliminate the mass variables, measuring lengths and times in terms of the total mass M and momenta in terms of the reduced mass μ . In these units, consistency then forces the angular momenta to be measured in terms of μM .

A second implementation uses a *Mathematica* interface as a front-end to an optimized C++ back-end. For long-time integrations, the increase in speed is worth the effort. Moreover, the calculation of Lyapunov exponents (Sec. 4.5.1 below) requires access to the internals of the ODE solver, so `NDSolve` is inappropriate. Instead, we use the techniques and routines described in Chapter 5. In order to calculate the Lyapunov exponents of any dynamical system, in general we need both the equations of motion and the Jacobian matrix of the system. We use *Mathematica* to generate the required expressions directly from the Hamiltonian, using the built-in `CForm` function to convert to C++ source code.⁴ We then link these routines to the general ODE integrator and Lyapunov calculator described in Chapter 5. *Mathematica*'s ability to handle pipes and temporary files makes it possible to use exactly the same interface for both the native *Mathematica* ODE solver and the C++ back-end, so that the difference between the two implementations is transparent to the user.

As noted in Sec. 4.2.2, we verify in all cases that the relevant quantities are conserved. Additionally, we find that, in the appropriate limits, the equations reproduce the Keplerian orbits of Newtonian gravitation (N turned on), the Lense-Thirring precession (extreme mass-ratio limit, N and SO turned on), and the (classical) quadrupole precession (N and $S_1 S_1$ turned on). Finally, the Lyapunov routines and Jacobian matrix used in Sec. 4.5 are verified using the techniques from Chapter 5.

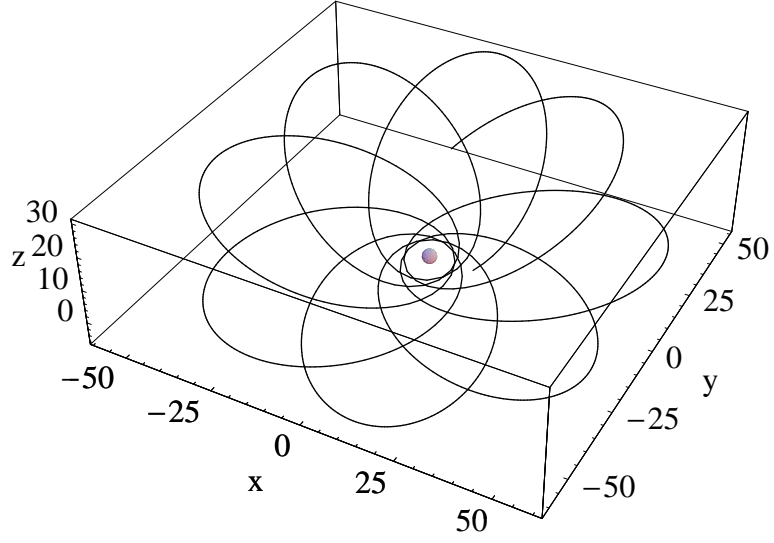
4.3 Parameterizing post-Newtonian orbits

We discuss here two convenient methods for parameterizing initial conditions for post-Newtonian orbits. We first describe a method that gives orbits that approximately satisfy desired values of eccentricity, pericenter, and orbital inclination. We then treat the important special case of quasi-circular orbits. Finally, we examine the effects of varying the post-Newtonian terms included in the initial conditions and subsequent time-evolution.

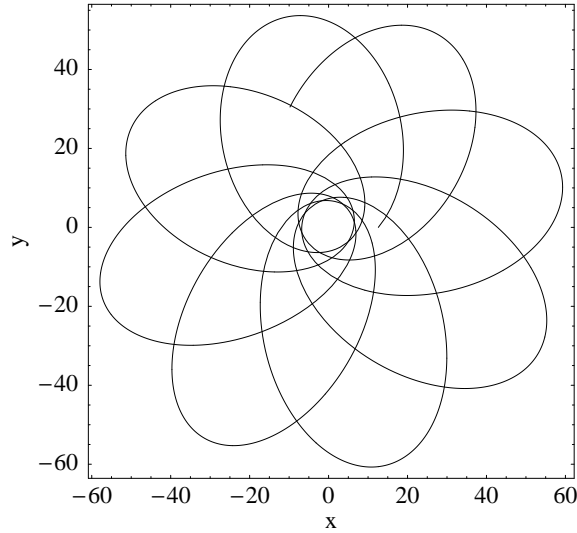
4.3.1 Eccentric orbits

One convenient parameterization method allows the easy creation of bound orbits and makes contact with the Kerr (geodesic) and Newtonian limits, namely, parameterizing in terms of the eccentricity e , pericenter r_p , and orbital inclination angle ι . These three quantities are obviously constants of the motion in the Newtonian limit, but in fact they can also be defined so that they correspond to conserved quantities for geodesics in Kerr spacetime (to which the PN equations approximately reduce in the limit where one of the masses dominates the other, say $m_2 \ll m_1$, and the system is not strongly relativistic, so that $v \ll 1$ and $r = |\mathbf{X}| \ll M$).

⁴The additional use of the freely available package `Optimize.m` gives a factor of 5 increase in speed for the case at hand.



(a)



(b)

Figure 4.2: A highly eccentric orbit for two maximally spinning $10 M_{\odot}$ black holes. (a) The orbit embedded in Euclidean space; (b) the projection onto the x - y plane. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$. The energy and angular momentum of the orbit correspond to a Kerr geodesic orbit with eccentricity $e = 0.8$, pericenter $r_p = 7 M$, and orbital inclination $\iota = 30^\circ$ ($E = 0.985915 \mu M$ and $L_z = 3.38238 \mu M$, where $\mu = m_1 m_2 / M$ is the reduced mass). The empirical values of these parameters, as determined from the numerical solution to the equations of motion, are $e = 0.805$, $r_p = 6.72 M$, and $\iota = 30.9^\circ$.

Given the desired values of the orbital parameters (e, r_p, ι) , and the masses and spins of the bodies, we must find valid initial conditions for the relative position \mathbf{X} and relative momentum \mathbf{P} . Our method relies on a mapping between (e, r_p, ι) and (E_K, L_z, Q) for families of Kerr geodesics, where E_K , L_z , and Q are the energy, z angular momentum, and Carter constant. (This mapping is described in detail in Chapter 3.) In the present case of the PN equations, we take in as parameters the requested values in the set (e, r_p, ι) , together with the masses and spins. We then calculate the appropriate value of the Kerr spin parameter a , which in our case is simply the magnitude of \mathbf{S}_{eff} [10]:

$$a = \|\mathbf{S}_{\text{eff}}\|. \quad (4.17)$$

Finally, we set $\phi = 0$, $\theta = \pi/2$, and $r = r_p(1 + e)$. Using these values of a , r , θ , and ϕ , we find the values of the set (E_K, L_z, Q) corresponding to (e, r_p, ι) for a Kerr geodesic, exactly as in Chapter 3.

These conserved quantities allow the calculation of the corresponding Kerr momentum [13], but this momentum does not correspond to the energy for the PN equations (as given by the PN Hamiltonian evaluated at the initial conditions). Instead, we keep two of the Kerr momentum components by setting $P_{x,\text{PN}} = p_{x,\text{Kerr}}$ and $P_{y,\text{PN}} = p_{y,\text{Kerr}} = L_z/r$. We then solve the equation

$$H(\mathbf{X}, \mathbf{P}, \mathbf{S}_1, \mathbf{S}_2) = E_K \quad (4.18)$$

for $P_{z,\text{PN}}$. The resulting momentum, when combined with the initial position and spins, gives a set of initial conditions for the PN equations with the same energy and z angular momentum as a Kerr geodesic with orbital parameters (e, r_p, ι) and spin parameter $a = S_{\text{eff}}$. The resulting orbits have empirical values of (e, r_p, ι) quite close to those requested, even for equal-mass binaries (Fig. 4.2).

4.3.2 Quasi-circular orbits

A second, more specialized parameterization of the PN initial conditions enforces the condition of *quasi circularity*. In particular, through second post-Newtonian order the quasi-circular orbits are in fact exactly circular, and even with spin-orbit coupling added there exist “spherical” orbits, i.e., orbits confined to lie on a sphere, with fixed radius but varying angle θ . Once any of the spin-spin terms is turned on, exact sphericity is impossible in general, but it is still possible to satisfy exactly the conditions leading to spherical orbits in the absence of spin-spin coupling. These orbits are especially important for modeling possible sources of gravitational radiation, since the orbits of compact binaries are expected to circularize due to gravitational radiation reaction, and then remain circular [14, 15].

The conditions leading to quasi-circular orbits are as follows. Given an initial radius r_0 , we set

$\phi_0 = 0$ and $\theta_0 = \pi/2$, so that

$$\mathbf{X}_0 = (r_0, 0, 0). \quad (4.19)$$

We then require that the initial radial momentum vanish:

$$(P_r)_0 = 0. \quad (4.20)$$

[Since the Hamiltonian is quadratic in P_r , this means that $(\dot{r})_0 = (\partial H / \partial P_r)_{P_r=0} = 0$, so that (at least initially) the radius is not changing.] Finally, we require that the initial values of \dot{P}_r and $\dot{\theta}$ vanish, which means (using Hamilton's equations) that

$$\left(\frac{dP_r}{dt} \right)_0 = - \left(\frac{\partial H}{\partial r} \right)_0 = 0 \quad (4.21)$$

and

$$\left(\frac{d\theta}{dt} \right)_0 = - \left(\frac{\partial H}{\partial P_\theta} \right)_0 = 0. \quad (4.22)$$

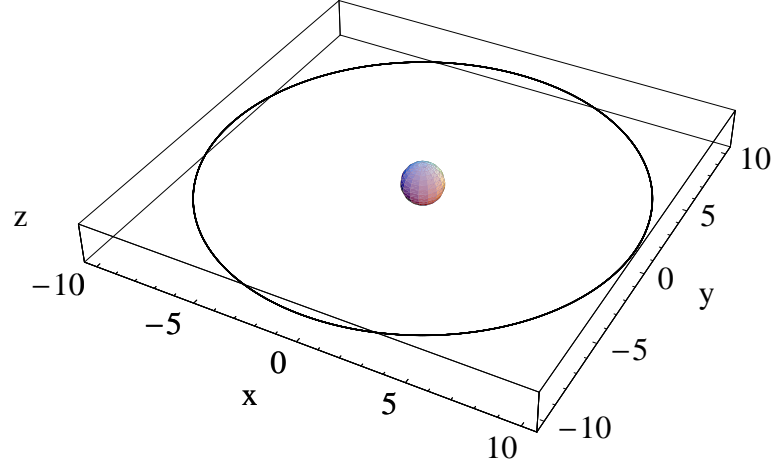
Given the initial position and the initial spins, these equations can be solved numerically for the initial values of P_θ and P_ϕ , thereby giving a complete set of initial conditions. One particular virtue of these conditions is that they reduce properly to circular orbits in the absence of the spin terms, but can still be satisfied by orbits with nonzero spin couplings.

4.3.3 Post-Newtonian orbits at various orders

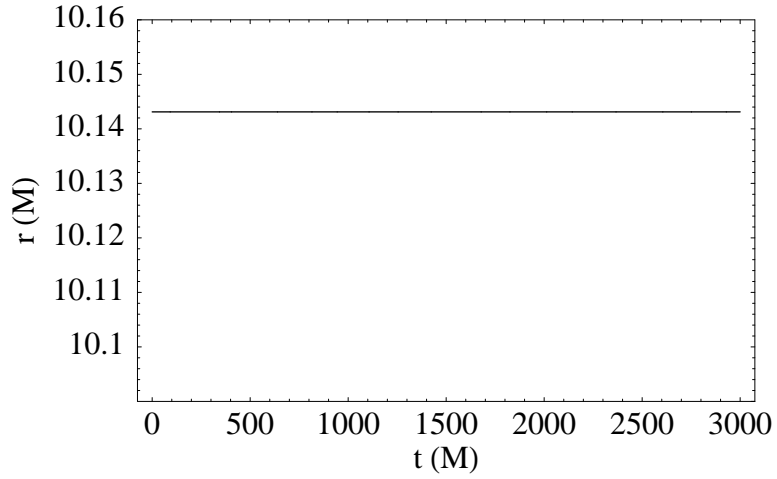
Here we show some of the effects of turning on or off the various post-Newtonian terms. Figs. 4.3–4.7 show quasi-circular orbits satisfying Eqs. (4.19)–(4.22), successively including more terms in the PN Hamiltonian from Sec. 4.2.1. We use quasi-circular orbits to make the effects especially easy to see, but this is not a necessary restriction. In particular, note that Figs. 4.1 and 4.2 demonstrate that the PN equations capture some of the essential aspects of relativistic orbits, especially the characteristic precession of the orbital plane.

4.4 The extreme mass-ratio case

When the masses satisfy $m_1 \gg m_2$, the PN equations describe a spinning test particle of mass m_2 orbiting a spinning supermassive black hole of mass m_1 . If we ignore the terms involving S_2 , then the resulting orbits are essentially geodesics orbiting a Kerr black hole, so by comparing these orbits to those with nonzero S_2 terms we can get a sense of the effects of the S_2 terms on the dynamics. Another motivation for considering the extreme mass-ratio case is to make contact with Chapters 2 and 3, which investigated the dynamics of spinning test particles orbiting Schwarzschild or Kerr



(a)



(b)

Figure 4.3: A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with Newtonian, 1PN, and 2PN terms turned on. (a) The orbit embedded in Euclidean space; (b) the radius $r = \sqrt{x^2 + y^2 + z^2}$ as a function of time. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$. The Newtonian gravitational-wave frequency is $f_{\text{GW}} = 100$ Hz, corresponding to a radius of $r = 10.14 M$ [Eq. (4.29)]. The initial spins are $\mathbf{S}_1 = (1, 0, 0) m_1^2$ and $\mathbf{S}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) m_2^2$, and the other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2).

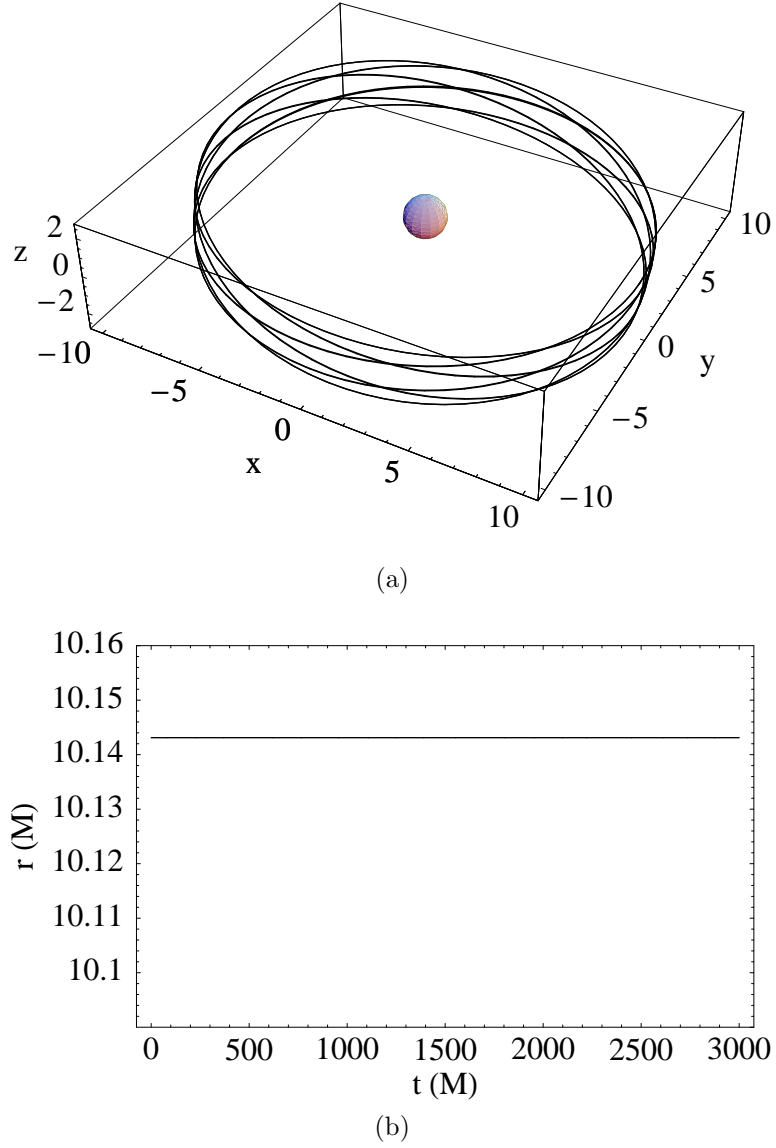


Figure 4.4: A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with terms through spin-orbit coupling. (a) The orbit embedded in Euclidean space; (b) the radius $r = \sqrt{x^2 + y^2 + z^2}$ as a function of time. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$. The Newtonian gravitational-wave frequency is $f_{\text{GW}} = 100$ Hz, corresponding to a radius of $r = 10.14 M$ [Eq. (4.29)]. The initial spins are $\mathbf{S}_1 = (1, 0, 0) m_1^2$ and $\mathbf{S}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) m_2^2$, and the other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2). The addition of spin-orbit coupling to the N, 1PN, and 2PN terms destroys exact circularity, but exact sphericity is preserved.

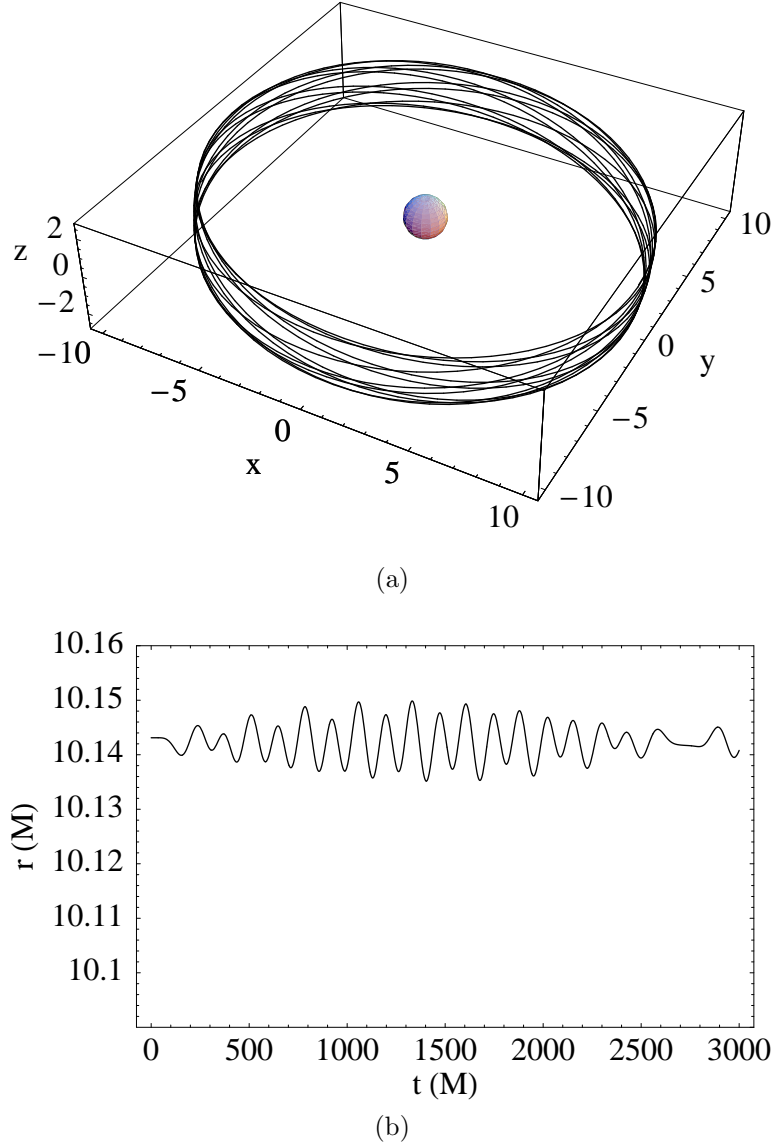


Figure 4.5: A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with all the terms from Sec 4.2.1 present except for the $S_1 S_1$ and $S_2 S_2$ spin-induced quadrupole terms. (a) The orbit embedded in Euclidean space; (b) the radius $r = \sqrt{x^2 + y^2 + z^2}$ as a function of time. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$. The Newtonian gravitational-wave frequency is $f_{\text{GW}} = 100$ Hz, corresponding to a radius of $r = 10.14 M$ [Eq. (4.29)]. The initial spins are $\mathbf{S}_1 = (1, 0, 0) m_1^2$ and $\mathbf{S}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) m_2^2$, and the other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2). The addition of spin-spin coupling to the N, 1PN, and 2PN terms destroys exact sphericity, but the Eqs. (4.19)–(4.22) are still satisfied, leading to nearly circular orbits for these initial conditions.

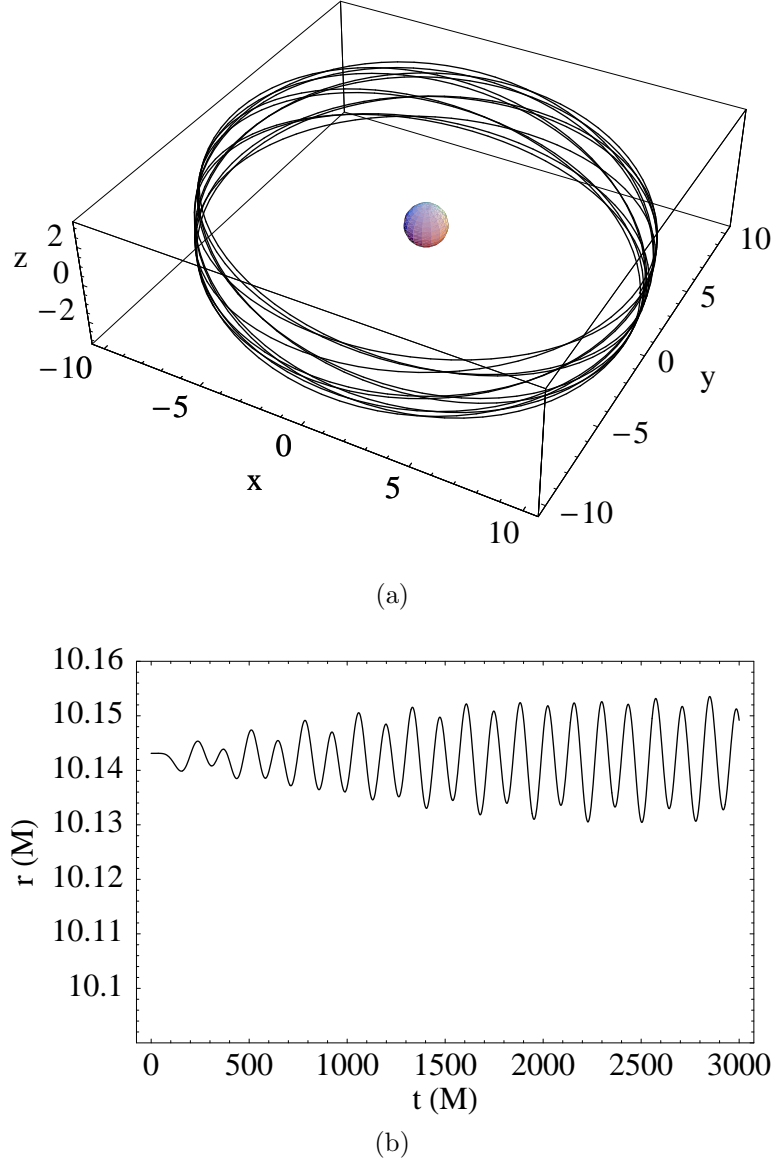


Figure 4.6: A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with all the terms from Sec 4.2.1 present except for the $S_2 S_2$ spin-induced quadrupole term. (a) The orbit embedded in Euclidean space; (b) the radius $r = \sqrt{x^2 + y^2 + z^2}$ as a function of time. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$. The Newtonian gravitational-wave frequency is $f_{\text{GW}} = 100$ Hz, corresponding to a radius of $r = 10.14 M$ [Eq. (4.29)]. The initial spins are $\mathbf{S}_1 = (1, 0, 0) m_1^2$ and $\mathbf{S}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) m_2^2$, and the other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2). The addition of spin-spin coupling to the N, 1PN, and 2PN terms destroys exact sphericity, but the Eqs. (4.19)–(4.22) are still satisfied, leading to nearly circular orbits for these initial conditions.

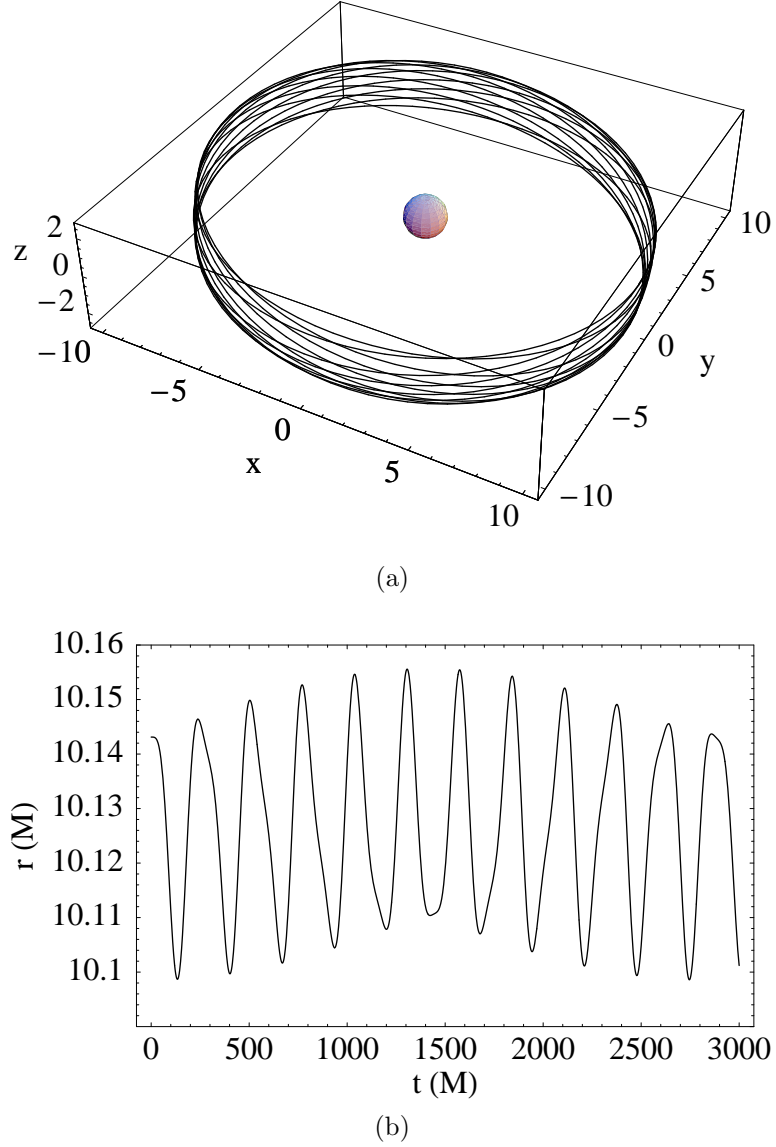


Figure 4.7: A post-Newtonian quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes, with all the terms from Sec 4.2.1 (including the all the spin-spin couplings) present. (a) The orbit embedded in Euclidean space; (b) the radius $r = \sqrt{x^2 + y^2 + z^2}$ as a function of time. Lengths are measured in terms of the total mass $M = m_1 + m_2$, and we show a schematic horizon at $r_H = M$. The Newtonian gravitational-wave frequency is $f_{\text{GW}} = 100$ Hz, corresponding to a radius of $r = 10.14 M$ [Eq. (4.29)]. The initial spins are $\mathbf{S}_1 = (1, 0, 0) m_1^2$ and $\mathbf{S}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) m_2^2$, and the other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2). The addition of spin-spin coupling to the N, 1PN, and 2PN terms destroys exact sphericity, but the Eqs. (4.19)–(4.22) are still satisfied, leading to nearly circular orbits for these initial conditions.

Table 4.1: Phase shifts $\Delta\phi = \phi_{\text{geodesic}} - \phi_{\text{spin}}$ in radians as a function of orbital inclination angle ι and pericenter r_p for $e = 0.5$ in the extreme mass-ratio case $m_2/m_1 = 10^{-5}$. (The entries scale roughly as m_2/m_1 .) The spin of the small body is offset $\pi/4$ toward the spin of the large central body, and both have maximal spins. The inclination angle indicates whether the orbit is prograde ($0 \leq \iota < 90^\circ$) or retrograde ($90^\circ \leq \iota \leq 180^\circ$). The “geodesic” orbits and their corresponding spin orbits start with the same initial 4-velocity \vec{v} , and the integrations use $t_{\text{max}} \approx (10 \text{ times the average radial orbital period})$. Blank entries indicate radius/inclination combinations for which no valid initial condition was found.

ι	$r_p(M)$				
	4	8	12	16	20
10°			4.8×10^{-6}	2.6×10^{-6}	1.3×10^{-6}
45°	-4.2×10^{-5}	-8.9×10^{-6}	-5.3×10^{-6}	-2.6×10^{-6}	-5.1×10^{-6}
85°	-1.6×10^{-4}	-1.5×10^{-4}	-1.9×10^{-5}	-9.5×10^{-6}	-3.0×10^{-6}
135°		6.9×10^{-5}	8.5×10^{-5}	2.9×10^{-5}	118×10^{-5}
170°		7.9×10^{-5}	4.4×10^{-5}	1.8×10^{-5}	9.4×10^{-6}

Table 4.2: Phase shifts $\Delta\phi = \phi_{\text{geodesic}} - \phi_{\text{spin}}$ in radians as a function of orbital inclination angle ι and pericenter r_p for quasi-circular orbits in the extreme mass-ratio case $m_2/m_1 = 10^{-5}$. (The entries scale roughly as m_2/m_1 .) The spin of the small body is offset $\pi/4$ toward the spin of the large central body, and both have maximal spins. The inclination angle indicates whether the orbit is prograde ($0 \leq \iota < 90^\circ$) or retrograde ($90^\circ \leq \iota \leq 180^\circ$). The “geodesic” orbits and their corresponding spin orbits start with the same initial 4-velocity \vec{v} , and the integrations use $t_{\text{max}} \approx (10 \text{ times the average } \phi \text{ orbital period})$.

ι	$r_p(M)$				
	4	8	12	16	20
10°	-5.8×10^{-5}	4.7×10^{-6}	5.3×10^{-6}	3.7×10^{-6}	3.2×10^{-6}
45°	1.6×10^{-4}	2.2×10^{-5}	1.2×10^{-3}	6.8×10^{-6}	4.6×10^{-6}
85°	2.3×10^{-5}	2.1×10^{-5}	-8.5×10^{-6}	-3.3×10^{-6}	-2.2×10^{-8}
135°	4.4×10^{-5}	-2.3×10^{-5}	-9.9×10^{-5}	-1.3×10^{-5}	-3.3×10^{-7}
170°	-9.9×10^{-5}	-4.2×10^{-5}	-1.9×10^{-5}	-9.6×10^{-6}	-5.6×10^{-6}

black holes using the Papapetrou equations.

4.4.1 Deviation from “geodesic” orbits

We can measure the effect of the S_2 terms in the equations of motion as follows. After choosing some appropriate final time t_f , we calculate the orbital phase angle ϕ_g for “geodesics,” which includes the PN terms through spin-orbit coupling, together with the $H_{S_1 S_1}$ term to account for the quadrupole moment of the central black hole. Then, for the same t_f , we calculate ϕ_s , the orbital phase angle with the S_2 terms ($H_{S_1 S_2}$ and $H_{S_2 S_2}$) turned on. The difference $\Delta\phi \equiv \phi_g - \phi_s$ provides a quantitative measure of the effect of the spin-spin coupling terms.

We present two tables of $\Delta\phi$ values, one for an eccentric orbit (Table 4.1, $e = 0.5$), and another for a quasi-circular orbit (Table 4.2, $e \approx 0$). The final time in both cases is ten times an effective

orbital period, defined as follows. For the eccentric case, we calculate the average radial period T_r , which we define as twice the average time between the turning points in the radial motion. In practice, this involves integrating for some time $t \gg T_r$, creating an interpolating function for r as a function of time, differentiating it, and then solving numerically for all the zeros.⁵ The average radial period is then twice the average time between zeros. Since the radius is nearly constant for the quasi-circular case, the use of T_r is inappropriate, and instead we use the average ϕ period T_ϕ , which is simply the final time divided by the winding number $\phi_f/2\pi$. Tables 4.1 and 4.2 show values of $\Delta\phi$ for varying orbital inclination angles and pericenters and a mass ratio of $m_2/m_1 = 10^{-5}$; the phase shifts scale roughly as m_2/m_1 , so these tables make it possible to obtain estimates for other mass ratios as well.

4.4.2 Comparing post-Newtonian and Papapetrou orbits

We mentioned in the introduction the investigations of the dynamics of a spinning test particle in a background spacetime. The *Papapetrou equations* model this system; if we choose the background to be the Kerr metric (as in Chapters 2 and 3), then the Papapetrou equations model the orbit of a small rotating black hole orbiting a supermassive rotating black hole. The Papapetrou system should therefore agree closely with the post-Newtonian equations when $m_1 \gg m_2$ and $r = |\mathbf{X}| \gg M$.

We can compare the PN equations directly with the Papapetrou equations (as reformulated by Dixon [16]) by examining the covariant time-derivative of the 4-momentum:

$$\frac{Dp^\mu}{d\tau} = -\frac{1}{2}R^\mu_{\nu\alpha\beta}v^\nu S^{\alpha\beta}, \quad (4.23)$$

where $R^\mu_{\nu\alpha\beta}$ is the Riemann tensor of the spacetime, v^ν is the 4-velocity, and $S^{\alpha\beta}$ is the spin tensor. For vanishing spin, this is simply the geodesic equation, so the Papapetrou equations include the Newtonian, 1PN, and 2PN terms. Since the Kerr metric is an exact description of a rotating black hole, the geodesic equation also includes the S_1 part of the spin-orbit coupling in Eq. (4.6), as well as the S_1S_1 quadrupole term [Eq. (4.10)]. [Here we identify S_1/m_1 with the Kerr parameter a , the spin angular momentum per unit mass. This is valid in the extreme mass-ratio limit, but see Eq. (4.17) above for the more general expression.]

For nonvanishing spin, the right-hand term in Eq. (4.23) is known to reproduce the post-Newtonian spin-orbit coupling to leading order when expanded in v^2/c^2 (that is, 1PN) [17]. This means that the spatial part of the expression $\frac{1}{2}R^\mu_{\nu\alpha\beta}v^\nu S^{\alpha\beta}$, expanded to lowest order in v^2/c^2 , is equal to the piece of $d\mathbf{P}/dt$ given by H_{SO} in Eq. (4.1). (There is an important subtlety here: the definitions for the center of mass used for the Papapetrou equations and the Hamiltonian PN equations differ by a factor of order 1PN: $\mathbf{X}_{\text{cm, PN}} = \mathbf{X}_{\text{cm, Papa}} + \frac{1}{2\mu}\mathbf{v} \times \mathbf{S}$. The correspondence between the

⁵The freely available *Mathematica* package `RootSearch` proved invaluable in this context.

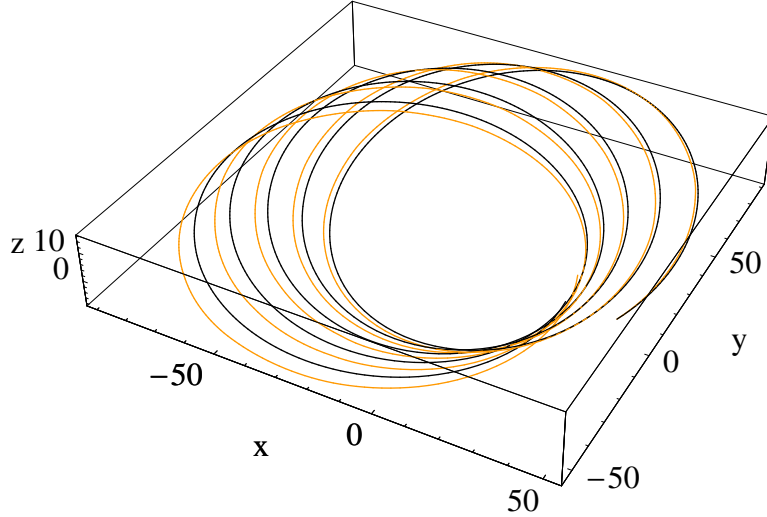


Figure 4.8: The orbit of two maximally spinning black holes in the extreme mass ratio ($m_1 \gg m_2$) limit, using the post-Newtonian equations (dark) and the Papapetrou equations (light). The Papapetrou initial conditions satisfy the conditions of eccentricity $e = 0.5$, pericenter $r_p = 30$, and orbital inclination $\iota = 10^\circ$. The analogous PN orbit starts with the same position and momentum. In both cases, the spin of the small body initially points 45° from the vertical. The PN equations of motion include all the PN terms from Sec. 4.2.1 except the $S_2 S_2$ quadrupole term.

Papapetrou spin-orbit coupling and the PN spin-orbit coupling is only evident once this substitution has been made.) To higher order, the right-hand side of Eq. (4.23) includes products of S with the Riemann tensor, which is a function of $a = S_1/m_1$; since S in the Papapetrou equations is S_2 in our treatment of the PN equations, this means that Eq. (4.23) should include the $S_1 S_2$ coupling term [Eq. (4.9)], although this has not to our knowledge ever been verified explicitly. Finally, the Papapetrou equations model a *pole-dipole* particle, so they do not include the $S_2 S_2$ quadrupole term of the smaller body [Eq. (4.11)].

We can compare Papapetrou and PN orbits more directly by constructing analogous initial conditions. We first select values of the eccentricity e , the pericenter r_p , and the orbital inclination angle ι for the Papapetrou equations, as discussed in Chapter 3. We then use the resulting values of position, momentum, and spin as the initial conditions for the PN equations of motion (with the $S_2 S_2$ term turned off). The orbits agree closely, especially at higher radii, as shown in Fig. 4.8. The quantitative agreement is worse at lower radii due to the violation of the condition $r \gg M$ required for the validity of the PN equations, but the qualitative behavior—especially the precession of the orbital plane—is essentially the same.

As a more quantitative comparison, we can use the PN equations to investigate the errors in the Papapetrou system caused by deviations from the test particle approximation. As discussed at length in Chapter 2, if we measure distances in terms of the central mass M_P and momenta in terms

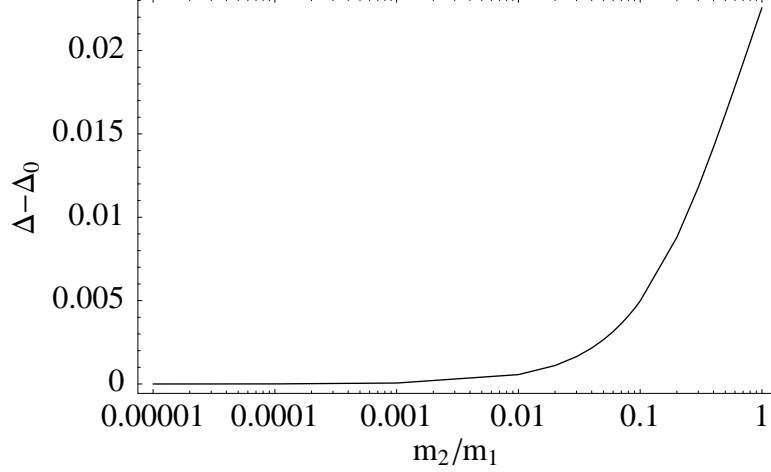


Figure 4.9: The error in the Papapetrou radial period $T_{r,P}$ vs. mass ratio for an eccentric ($e = 0.5$) orbit with pericenter $r_P = 30$. Here Δ is the fractional difference between the Papapetrou radial period and the post-Newtonian radial period: $\Delta = T_{r,P}/T_{r,PN} - 1$. Even in the $m_2 \rightarrow 0$ limit, we are unable to find initial conditions leading to exactly the same periods, which we correct by subtracting the fractional difference Δ_0 , which is the extrapolation of Δ to the case of vanishing mass ratio. (In this case, $\Delta_0 = 0.003$.) It is evident that the Papapetrou radial period differs significantly from the corresponding PN period when $m_2 \gtrsim 0.1 m_1$, indicating a failure of the test-particle approximation.

of the particle mass μ_P , then the spin S of the small body is measured in terms of $\mu_P M_P$. (Here we use the subscript P for Papapetrou, to distinguish the Papapetrou masses from the total mass M and reduced mass μ used for the PN system.) This means that a small maximally spinning black hole in a maximal Kerr background has a spin parameter of

$$S = \frac{\text{maximum spin}}{\mu_P M_P} = \frac{\mu_P^2}{\mu_P M_P} = \frac{\mu_P}{M_P} \ll 1. \quad (4.24)$$

The last step uses the requirement that the second body be a test particle. Deviations from this approximation are equivalent to increasing the value of S , so that $S = 1$ corresponds to $\mu_P = M_P$, i.e., an equal mass binary.

Our strategy is to consider initial conditions where both the Papapetrou system and PN system are valid: orbits with extreme mass ratios (Papapetrou valid) and relatively large pericenters (PN valid). Using the average radial period T_r , we can compare the two systems quantitatively as we vary the mass ratio: from its initial small value (taken to be 10^{-5}) we increase m_2 until the mass ratio is 1. Since the PN equations are valid in the equal mass ($m_1 = m_2$) limit, the difference between the Papapetrou and PN values of T_r is due to the violation of the test particle approximation. The results appear in Fig. 4.9 for an eccentric orbit and Fig. 4.10 for a near-circular orbit. In both cases, the Papapetrou period deviates significantly from the PN value when $m_2 \gtrsim 0.1 m_1$, which indicates a failure of the test-particle approximation.

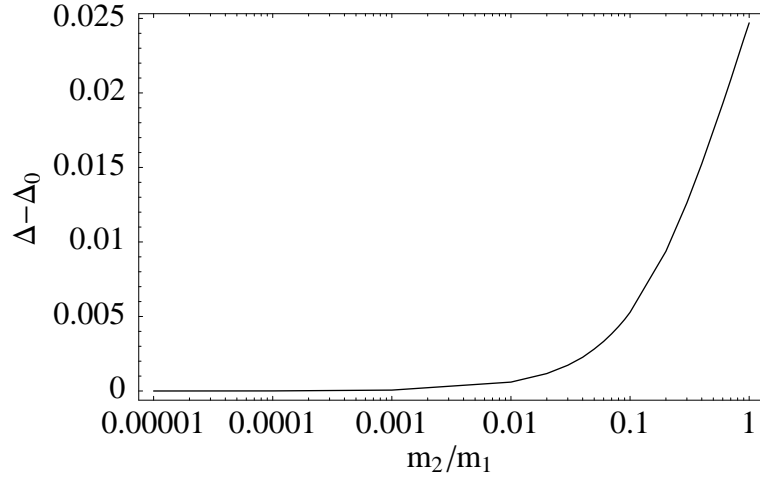


Figure 4.10: The error in the Papapetrou radial period $T_{r,P}$ vs. mass ratio for a near-circular ($e = 0.05$) orbit with pericenter $r_P = 30$. As in Fig. 4.9, Δ is the fractional difference between the Papapetrou radial period and the post-Newtonian radial period ($\Delta = T_{r,P}/T_{r,PN} - 1$), and Δ_0 is the extrapolation of Δ to the case of vanishing mass ratio. (In this case, $\Delta_0 = 0.001$.) The Papapetrou radial period differs significantly from the corresponding PN period when $m_2 \gtrsim 0.1 m_1$, indicating a failure of the test-particle approximation.

4.5 Investigating chaos in the post-Newtonian equations

Previous studies of chaos in the post-Newtonian equations considered comparable mass-ratio binaries with eccentric orbits [2, 4, 19–21]. Here, we also consider comparable mass-ratio binaries, but choose to focus on quasi-circular orbits. As noted in the introduction, these orbits are particularly important because many astrophysically relevant binary systems should circularize due to the energy lost to gravitational radiation. We set the radii of these orbits so that the frequencies of their corresponding gravitational waves lie in a range $40 \text{ Hz} \leq f_{\text{GW}} \leq 240 \text{ Hz}$, roughly corresponding to the frequency band for the Laser Interferometer Gravitational-Wave Observatory (LIGO). [More properly, we choose radii such that the *Newtonian* frequency of the gravitational waves (which is simply twice the orbital frequency) lies in the LIGO band, as discussed in Sec. 4.5.2 below.]

4.5.1 A brief discussion of Lyapunov exponents

As in Chapters 2 and 3, Lyapunov exponents are our primary tool for investigating the nonlinear dynamics of general relativistic systems. We have discussed at length in Chapters 2 and 5 our techniques for calculating these exponents for systems similar to the PN equations. Here we present a brief summary of Lyapunov exponents.

Given an initial condition in the phase space of a dynamical system with n degrees of freedom, we imagine an n -dimensional ball of nearby initial conditions centered on that point. As the dynamics unfold, in general the ball is stretched in some directions and squeezed in others, deforming into an n -

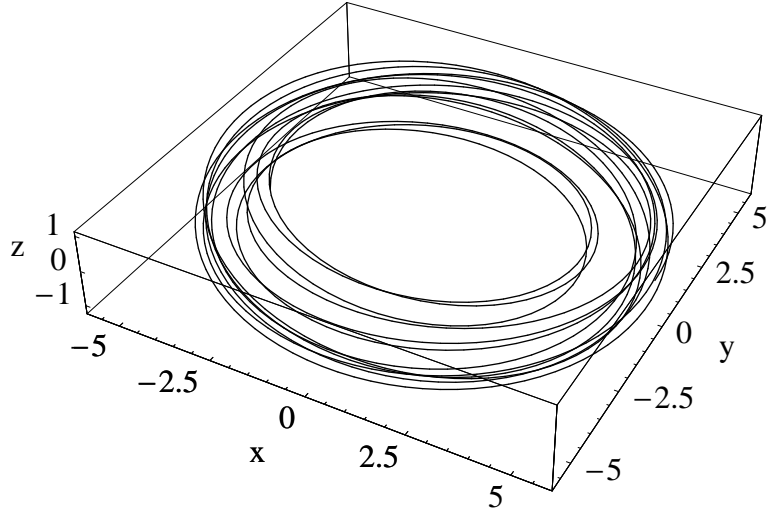


Figure 4.11: The orbit of two maximally spinning $10 M_{\odot}$ black holes. The dynamics are chaotic, as shown in Fig. 4.12. The initial conditions satisfy the requirements for quasi circularity, though in fact the orbit's radius is not even approximately constant (see Sec. 4.5.2.2 below). The initial radius is $5.658 M$, corresponding to an orbit with a Newtonian gravitational-wave frequency of $f_{\text{GW}} = 240 \text{ Hz}$ [Eq. (4.29)]. The initial spins are $\mathbf{S}_1 = (0.13036, 0.262852, -0.955989) m_1^2$ and $\mathbf{S}_2 = (0.118966, -0.13459, -0.983734) m_2^2$. The other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2).

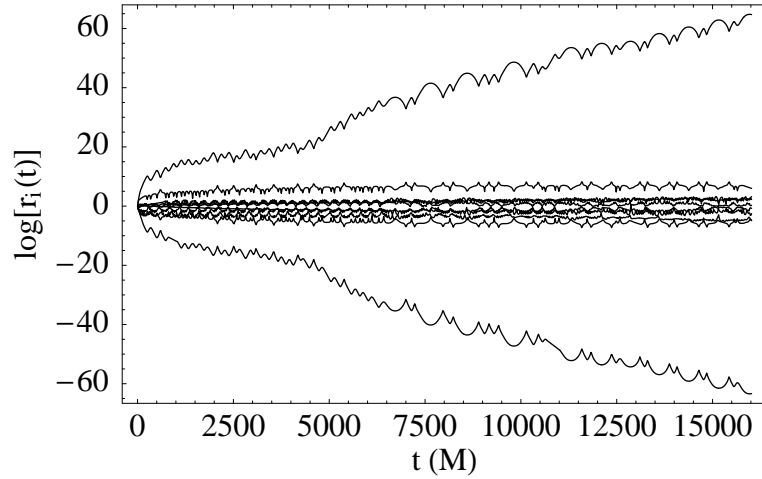


Figure 4.12: The natural logarithms of the ellipsoid axes r_i vs. t for the system shown in Fig. 4.11. The slopes of the lines are the Lyapunov exponents. Two nonzero exponents are clearly visible ($\lambda = \pm 3.2 \times 10^{-3} M^{-1}$), but all the others are consistent with zero. There is an apparent $\pm\lambda$ symmetry: for each exponent $+\lambda$, there is a corresponding exponent $-\lambda$; even the zero exponents approach zero symmetrically. This behavior is a characteristic of Hamiltonian systems [18].

dimensional ellipsoid under the action of the flow. Such an ellipsoid has n principal (semi)axes $r_i(t)$, and the average multiplicative rate of stretching or squeezing of each axis is the i th *Lyapunov number*:

$$L_i = \lim_{t \rightarrow \infty} [r_i(t)]^{1/t}. \quad (4.25)$$

The natural logarithm of the Lyapunov number is the *Lyapunov exponent* associated with the axis:

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{\log [r_i(t)]}{t}. \quad (4.26)$$

Implementing this prescription numerically leads to a visualization of the exponents as a plot of $\log [r_i(t)]$ vs. t (so that the slope is the exponent λ_i), which refer to as a *Lyapunov plot*. The result for a chaotic PN orbit appears in Fig. 4.12.

In practice, following the evolution of the phase-space ellipsoid, and thereby extracting all the Lyapunov exponents of the system, involves using the Jacobian matrix of the system to model an “infinitesimal” ball that captures the true linear approximation to the dynamics. It is also possible, and computationally faster, to extract only the largest Lyapunov exponent by considering only one nearby initial condition, joined by some small deviation vector to the original point, and the tracking the growth of the deviation with time. In what follows, most of our simulations use this faster (but less robust) *deviation vector* method, but we have checked many of the results using the Jacobian method. Further details of our various techniques for calculating Lyapunov exponents appear in Chapter 2 and especially Chapter 5.

It is worth noting that we can think of the PN system as *constrained*, since we wish to think of the spin magnitudes as fixed. In other words, given an initial spin vector, a “nearby” initial spin should point in a different direction but have the same magnitude. The system thus has only 10 true degrees of freedom (three for relative position and momentum, and two for each spin), and should therefore have only 10 Lyapunov exponents. The constraints lead to significant complications in calculating the Lyapunov exponents; we discuss in Chapter 5 several methods for addressing these complications.

The principal value of the largest Lyapunov exponent is that it provides the e -folding timescale $t_\lambda = 1/\lambda$ for the divergence of nearby trajectories. The formal definition of λ in Eq. (4.26) requires an infinite-time limit, but of course any numerical method for λ must introduce some finite cutoff. As a result, in general it is impossible to say with any certainty that a system is not chaotic—even if it appears that $\lambda_{\max} \rightarrow 0$ for some t_{cutoff} , chaos may yet manifest itself on longer timescales. Nevertheless, it is possible to calculate nonchaotic *baseline orbits* (corresponding, for example, to the PN terms through the spin-orbit coupling), whose nearby initial conditions still exhibit some nonzero (power-law) separation. If a suspected chaotic orbit has a Lyapunov exponent with a magnitude similar to a baseline orbit, it is “indistinguishable from zero” and is probably not chaotic.

For the present problem, the most relevant timescale is the inspiral time due to the energy loss from gravitational radiation. For the quasi-circular orbits considered below, this is approximately given by the following formula [22], which is valid for circular orbits:

$$t_{\text{insp}} = \frac{5}{256} \frac{M^2}{\mu} \left(\frac{r}{M} \right)^4, \quad (4.27)$$

M is the total mass and μ is the reduced mass. We then adopt the criterion $t_\lambda < t_{\text{insp}}$ as an operational definition of chaos, which is equivalent to the condition

$$\lambda t_{\text{insp}} > 1 \quad \text{condition for chaotic orbits.} \quad (4.28)$$

On the other hand, if $t_\lambda > t_{\text{insp}}$, then, even if the system is formally chaotic in the conservative limit we consider here, the chaos will not have time to manifest itself before the final plunge. If there is no exponential divergence in a time t_{insp} , there is no significant difference in the number of templates needed for chaotic versus nonchaotic orbits, so in these cases chaos is unimportant for gravitational wave detection.

4.5.2 A survey of quasi-circular orbits

In this section, we elucidate the effects of varying the parameters in the PN equations of motion on the presence of chaos in the resulting dynamics. In many of the examples, we parameterize the orbits by their radii, or, equivalently, by the “gravitational wave frequency”:

$$f_{\text{GW}} = \frac{1}{\pi} \left(\frac{GM}{r^3} \right)^{1/2}, \quad (4.29)$$

where we restore the factor of G so that the result is in Hz. It is essential to note that Eq. (4.29) is the *Newtonian* gravitational wave frequency, which is valid only for radii that satisfy $r \gg M$. Nevertheless, Eq. (4.29) provides a convenient way to parameterize the initial conditions by radius in a way that has transparent physical significance in the nonrelativistic limit. When we refer below to an orbit with gravitational wave frequency of (say) 240 Hz, we mean an orbit with a radius that satisfies Eq. (4.29) when $f_{\text{GW}} = 240$ Hz. It is important to remember that this is close to, but not exactly, the true frequency of the gravitational wave.

4.5.2.1 Varying spin directions

We illustrate the effect of varying the spin directions by generating a large number of quasi-circular orbits with randomly oriented (maximal) spins. For each spin configuration, we choose the radius corresponding to a gravitational wave with frequency $f_{\text{GW}} = 240$ Hz (the high end of the LIGO

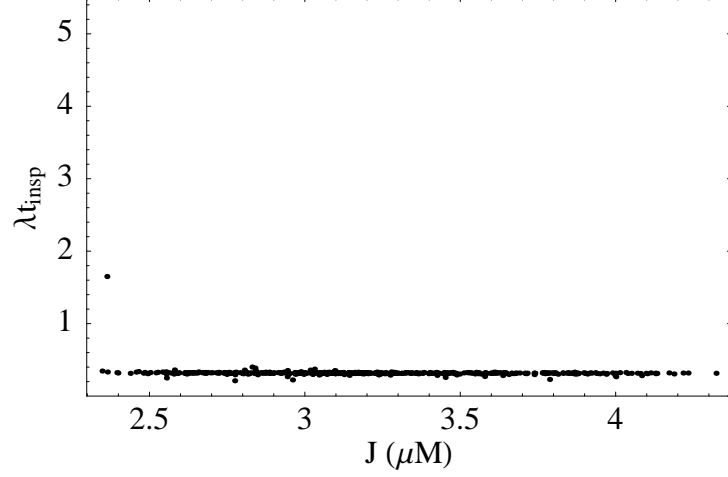


Figure 4.13: Lyapunov exponents for 500 quasi-circular orbits as a function of total angular momentum J for the $(10+10)M_{\odot}$ configuration. The spin for each body is maximal and randomly oriented (distributed uniformly on a unit sphere), and the initial radius corresponds to a gravitational wave frequency of $f_{\text{GW}} = 240$ Hz. The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$, so that $\lambda t_{\text{insp}} > 1$ indicates that nearby trajectories diverge by a factor of e on a timescale shorter than the inspiral timescale. There is only one such chaotic initial condition out of the 500 orbits considered for this configuration.

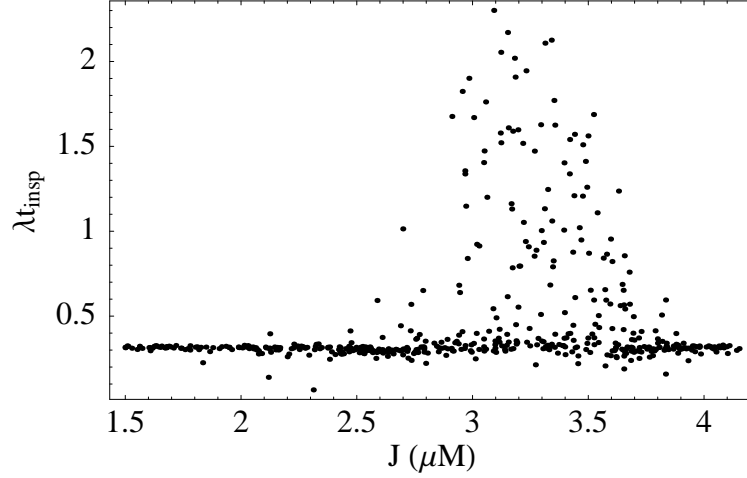


Figure 4.14: Lyapunov exponents for 500 quasi-circular orbits as a function of total angular momentum J for the $(20+10)M_{\odot}$ configuration at 240 Hz. The spin for each body is maximal and randomly oriented (distributed uniformly on a unit sphere), and the initial radius corresponds to a gravitational wave frequency of $f_{\text{GW}} = 240$ Hz. The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$, so that $\lambda t_{\text{insp}} > 1$ indicates that nearby trajectories diverge by a factor of e on a timescale shorter than the inspiral timescale. There are 49 such chaotic initial conditions out of the 500 orbits considered for this configuration. Unlike the $(10+10)M_{\odot}$ case shown in Fig. 4.15 below, the chaotic orbits in this case correspond to total angular momentum in the middle of the range.

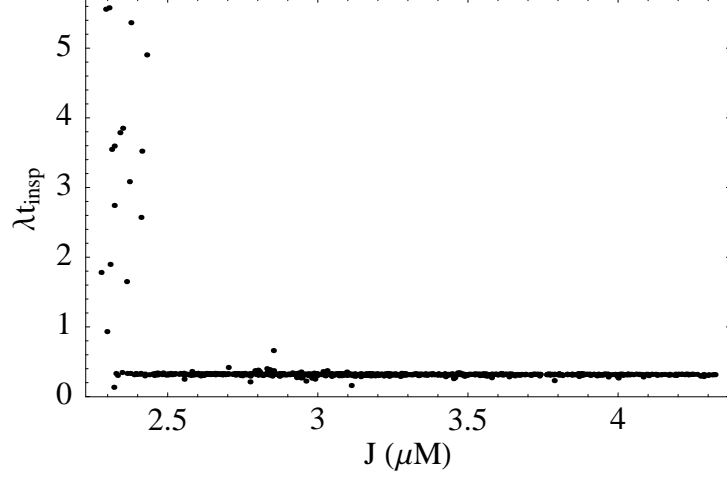


Figure 4.15: Lyapunov exponents for 500 quasi-circular orbits as a function of total angular momentum J for the $(10+10)M_{\odot}$ configuration at 240 Hz. The spin for each body is maximal with random initial spin angles (which overweights the poles); see Fig. 4.13 above for the results with the initial spin directions distributed randomly on a unit sphere. The initial radius corresponds to a gravitational wave frequency of $f_{\text{GW}} = 240$ Hz. The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$, so that $\lambda t_{\text{insp}} > 1$ indicates that nearby trajectories diverge by a factor of e on a timescale shorter than the inspiral timescale. There are 14 such chaotic initial conditions out of the 500 orbits considered for this configuration. They are clustered at the low end of the angular momentum range, indicating that the spins are closely aligned with each other and anti-aligned with the orbital angular momentum.

frequency band). The choice of radius is motivated by two main factors. First, choosing the lowest possible radius (consistent with the abilities to detect the corresponding gravitational waves) likely represents a worst-case scenario for chaos, since low-radius regions correspond to stronger spin couplings in the equations of motion (as noted in Chapter 3). Second, minimizing the radius minimizes the inspiral timescale t_{insp} , which in turn minimizes the computational cost of a final integration time significantly longer than t_{insp} . This allows us to achieve a better bound on Lyapunov exponents suspected to be zero, and increases our confidence that apparently nonzero exponents represent genuine chaotic behavior. In what follows, the final integration time is 10 times the inspiral time of each orbit.

We consider the following mass configurations: $(20+5)M_{\odot}$, $(10+5)M_{\odot}$, $(5+5)M_{\odot}$, $(10+10)M_{\odot}$, $(20+20)M_{\odot}$, $(20+10)M_{\odot}$, and $(15+5)M_{\odot}$. For each configuration we choose 500 initial conditions, with randomly oriented maximal initial spins whose directions are distributed uniformly on a unit sphere.⁶ The results appear in Figs. 4.13–4.15 and Table 4.3. The most notable result is the presence of chaotic orbits for the $(10+10)M_{\odot}$ and $(20+10)M_{\odot}$ cases. Fig. 4.16 shows a Lyapunov plot for the initial conditions producing the largest value λt_{insp} in our simulation data. The onset of chaos is marked by a transition from linear (or at most power-law) separation of nearby initial conditions

⁶One of several equivalent methods for achieving this is to choose u and v to lie uniformly in $[0, 1]$, and then choose $\theta = \cos^{-1}(2u - 1)$ and $\phi = 2\pi v$ for the initial spin orientation angles for each body.

Table 4.3: The prevalence of chaos in post-Newtonian quasi-circular orbits at 240 Hz, for spin directions chosen randomly on a unit sphere. We calculate the fraction of orbits whose e -folding times $t_\lambda = 1/\lambda$ are less than the inspiral time t_{insp} , which is our operational definition of chaos. The final integration time is 10 times the inspiral time. We also include 95% confidence intervals for the reported fractions, and we show the average value of λ measured in units of the inverse inspiral time for the $(20+10)M_\odot$ configuration (the only case in our simulation with more than one chaotic orbit). The simulation data represent 500 randomly chosen initial spin directions for each configuration, with the initial radius fixed by requiring a gravitational wave frequency of 240 Hz (as determined by the Newtonian formula).

Configuration	Fraction chaotic	95% confidence interval	Average chaotic λt_{insp}
$(20+5)M_\odot$	0	[0, 0.00738]	1.45
$(10+5)M_\odot$	0	[0, 0.00738]	
$(5+5)M_\odot$	0	[0, 0.00738]	
$(10+10)M_\odot$	0.002	$[5.06 \times 10^{-5}, 0.0111]$	
$(20+20)M_\odot$	0	[0, 0.00738]	
$(20+10)M_\odot$	0.104	[0.0777, 0.136]	
$(15+5)M_\odot$	0	[0, 0.00738]	

to exponential separation. On our Lyapunov plot (which is logarithmic on its vertical axis), chaotic orbits appear as linear growth.

Our initial simulation runs involved choosing random values of the spin angles, distributed uniformly in the appropriate ranges (i.e., $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$). This prescription does *not* result in points distributed uniformly on a unit sphere, but rather overweights points near the poles. In our case this was a lucky accident: as shown in Fig. 4.15, for $(10+10)M_\odot$ the chaotic orbits are clustered at the lowest values of the total angular momentum J , corresponding to initial spin vectors nearly anti-aligned with the orbital angular momentum \mathbf{L} , so that $J = |\mathbf{L} + \mathbf{S}|$ is minimized. This clustering is not visible in Fig. 4.13, which uses random points falling uniformly on a unit sphere; only one point out of 500 is chaotic in this case. The overweighting of the poles in Fig. 4.15 made possible the discovery that chaos in the $(10+10)M_\odot$ configuration is concentrated at values of the spin angles θ_1 and θ_2 near π , as shown in Fig. 4.17.

The clustering of chaos at low values of the total angular momentum in the $(10+10)M_\odot$ configuration suggests an explanation based on [23], which identifies a tumbling mode in spinning binaries due to the “loss of gyroscopic bearings” when the anti-alignment of \mathbf{S} and \mathbf{L} leads to a near-cancellation in their sum. While such a tumbling mode certainly seems a promising source of chaotic orbits, in our case the total angular momentum J is not much smaller than L or S (the condition for tumbling identified by [23]): for the chaotic orbits in Fig. 4.15, we have $J \approx 2.3\mu M$, $L \approx 4\mu M$, and $S \approx 2\mu M$. Moreover, the clustering at low total angular momentum is not a general result. In fact, when using uniformly distributed points, only the $(20+10)M_\odot$ configuration shows significant chaos, and its chaotic orbits occur mainly in the middle part of the angular-momentum

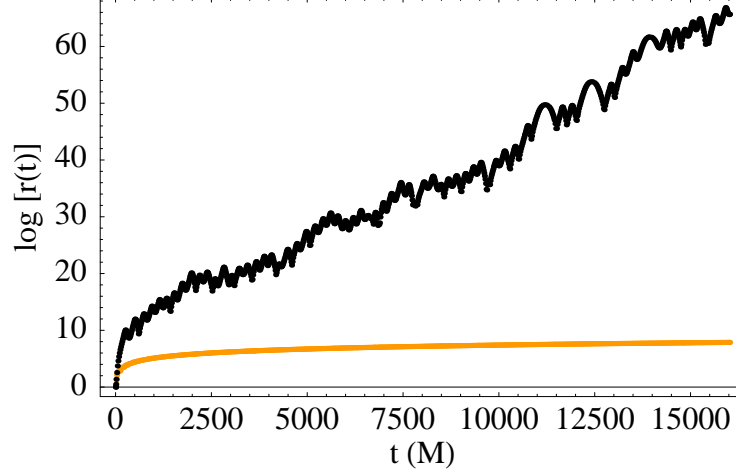


Figure 4.16: The natural logarithm of the principal ellipsoid axis vs. time for the strongest chaos (corresponding to the maximum value of λt_{insp}) found in the $(10 + 10)M_{\odot}$ case. The slope of the line is the Lyapunov exponent, which is approximately $\lambda = 3.2 \times 10^{-3} M^{-1}$ using a least-squares fit. This corresponds to a Lyapunov (e -folding) timescale of $t_{\lambda} = 1/\lambda = 3.1 \times 10^2 M$, which is less than a fifth of the inspiral timescale. The simulation data for a nonchaotic orbit (light) is shown for reference. The initial conditions producing this plot are the same as those in Fig. 4.11.

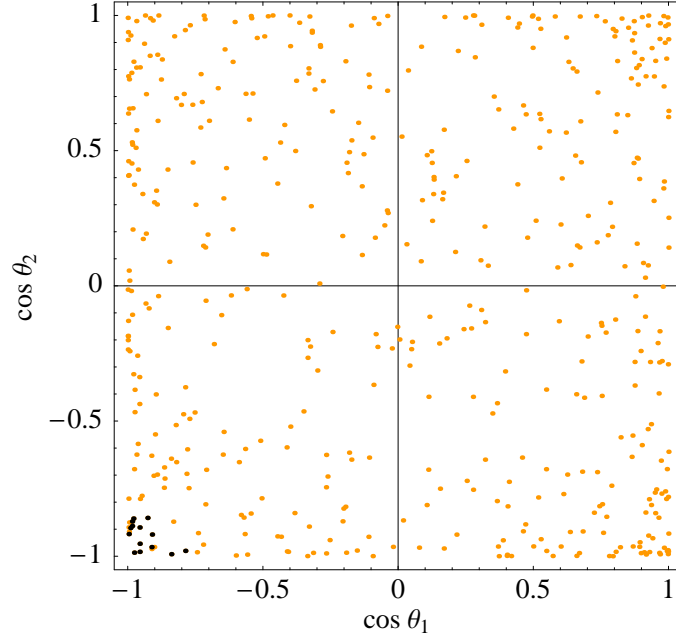


Figure 4.17: Chaos as a function of initial spin angle θ : $\cos \theta_2$ vs. $\cos \theta_1$ for the initial conditions in Fig. 4.15. Most points (light) are not chaotic; the few chaotic points (dark) are clustered in the lower left-hand corner, indicating that chaos [in this $(10 + 10)M_{\odot}$ case] occurs when the spins are nearly anti-aligned with the orbital angular momentum \mathbf{L} .

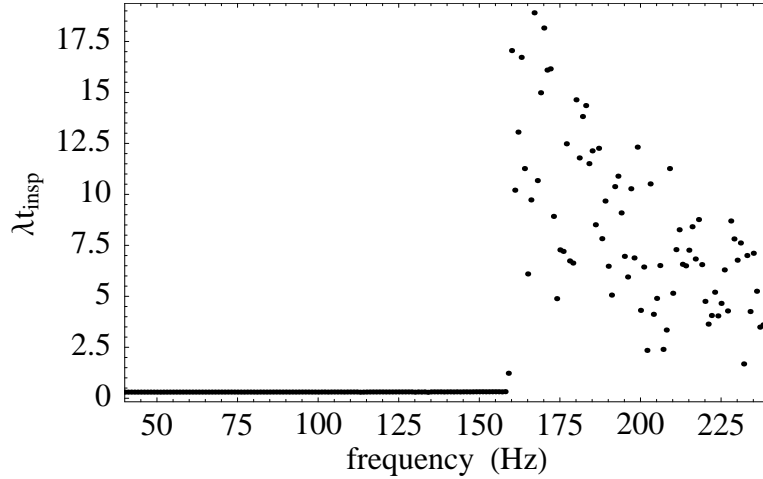


Figure 4.18: Lyapunov exponents as a function of gravitational-wave frequency for quasi-circular orbits of two maximally spinning $10 M_{\odot}$ black holes. The initial spins are the same as in Fig. 4.11, while the initial radius is determined by Eq. (4.29). The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$, and the frequencies are chosen to correspond closely to the LIGO frequency band. There is an abrupt transition to chaos at approximately 160 Hz. The subsequent decrease in λt_{insp} is purely the result of the decrease of t_{insp} with increasing frequency; in fact, the values of λ are approximately constant.

range (Fig. 4.14).

4.5.2.2 Varying initial frequencies

We now investigate the results of varying the initial gravitational-wave frequencies for the chaotic orbit from the previous section. We begin with a $(10 + 10)M_{\odot}$ system that is strongly chaotic when $f_{\text{GW}} = 240$ Hz (Fig. 4.11), and then vary the (Newtonian) frequency from 40 Hz back up to 240 Hz. (The initial radius is determined by the frequency, the initial spins are as given in Fig. 4.11, and the other initial conditions are determined by the conditions for quasi circularity.) The result appears in Fig. 4.18, which shows that chaos is absent for initial radii corresponding to f_{GW} less than 160 Hz. Above 160 Hz, there is an abrupt change in the dynamics from regular to chaotic, with a maximum Lyapunov exponent more than 18 times the inverse inspiral time (meaning that nearby trajectories diverge by a factor of e in a time $t_{\lambda} \approx t_{\text{insp}}/18$). Fig. 4.19 shows a Lyapunov plot for initial conditions on either side of the transition to chaos.

This qualitative change in the dynamical behavior in this case is mirrored in the orbits themselves. In particular, the onset of chaos is associated with a breakdown in the quasi circularity of the orbit. As shown in Fig. 4.20, just below the transition to chaos the orbit is nearly circular. Just above the transition, the orbits are not even approximately circular, as shown in Fig. 4.21.

A similar plot of λt_{insp} vs. frequency for the $(20 + 10)M_{\odot}$ appears in Fig. 4.22. There are a couple of scattered chaotic initial conditions at lower frequencies, but there is still a reasonably

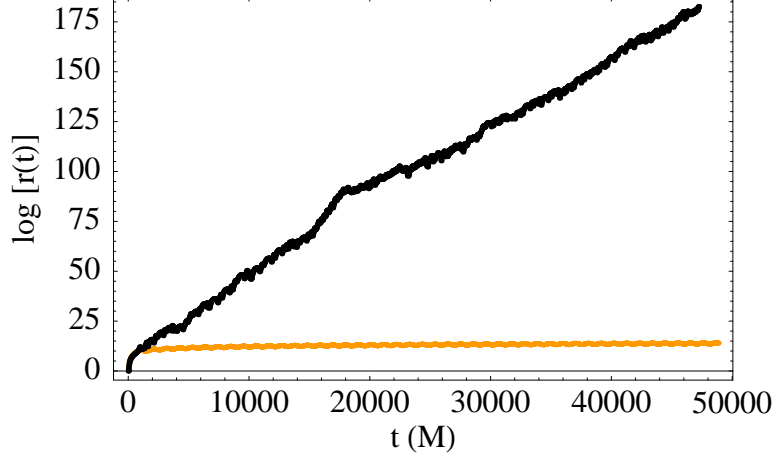


Figure 4.19: The natural logarithms of the principal ellipsoid axis vs. time for frequencies on opposite sides of the transition to chaos shown in Fig. 4.18: $f_{\text{GW}} = 158$ Hz (light) and $f_{\text{GW}} = 160$ Hz (dark). The slope of each line is the Lyapunov exponent, with $\lambda \approx 7 \times 10^{-5} M^{-1}$ (light, nonchaotic/consistent with zero) and $\lambda \approx 4 \times 10^{-3} M^{-1}$ (dark, chaotic). The initial spins are the same as in Fig. 4.11, while the initial radii are determined by Eq. (4.29). The orbits corresponding to these two frequencies appear in Figs. 4.20 and 4.21.

sharp transition to chaos at around 225 Hz, and above 230 Hz all the initial conditions (for this choice of initial spins) are chaotic. A Lyapunov plot of points before and after the transition appears in Fig. 4.23. The change in the orbits on either side of the transition is not as dramatic as in the the $(10 + 10)M_{\odot}$ system, but the orbit after the transition is noticeably more erratic, as shown in Figs. 4.24 and 4.25.

To verify that the disappearance of chaos at lower frequencies is generic, we repeated the 240 Hz survey for $f_{\text{GW}} = 40$ Hz in the $(10 + 10)M_{\odot}$ system. The inspiral times are very long in this case, requiring patience on the part of the simulator, but the results are gratifying: as shown in Table 4.4, we found not even one orbit with a Lyapunov time less than the inspiral time at 40 Hz. Any chaos, if present, manifests itself in this case on timescales longer than t_{insp} .

4.5.2.3 Varying spin magnitudes

In Sec. 4.5.2.2, we created one-parameter families of orbits by considering quasi-circular orbits that are chaotic at 240 Hz, and we then varied the frequency (or, equivalently, the quasi-circular radius) to determine the frequency dependence of chaotic behavior. In this section, we do the same, but instead fix the frequency (at 240 Hz) and vary the magnitude of one of the spins.

We begin with the maximally spinning $(10 + 10)M_{\odot}$ and $(20 + 10)M_{\odot}$ systems from Sec. 4.5.2.2 at 240 Hz, both of which are chaotic. We then vary the spin S_2 of the second body. As shown in Figs. 4.26 and 4.27, the dynamics are nonchaotic for most values of S_2 , with a transition to chaos at sufficiently high spins. The transition occurs at $S_2 = 0.85$ for the $(10 + 10)M_{\odot}$ configuration, and at

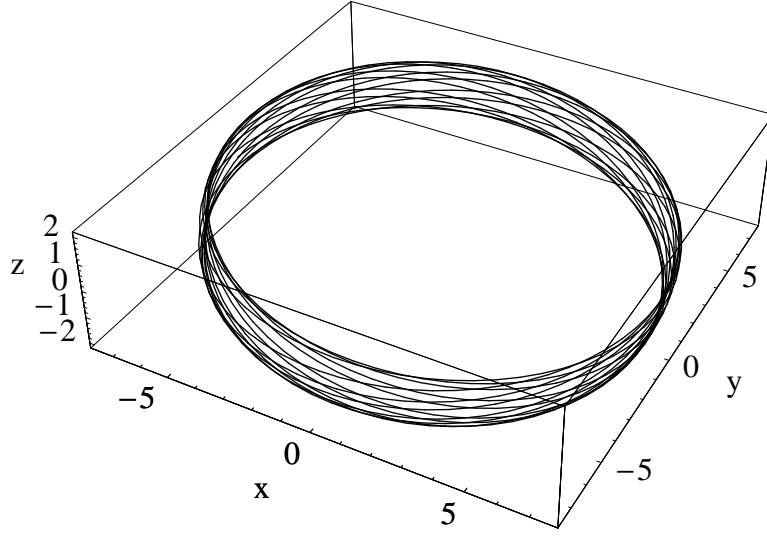


Figure 4.20: The nonchaotic quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes (in the effective one body approach), corresponding to a gravitational wave frequency of $f_{\text{GW}} = 158$ Hz. The initial spins are the same as in Fig. 4.11, while the initial radius is determined by Eq. (4.29). The corresponding Lyapunov plot is shown in Fig. 4.19. The orbit's radius is approximately constant.

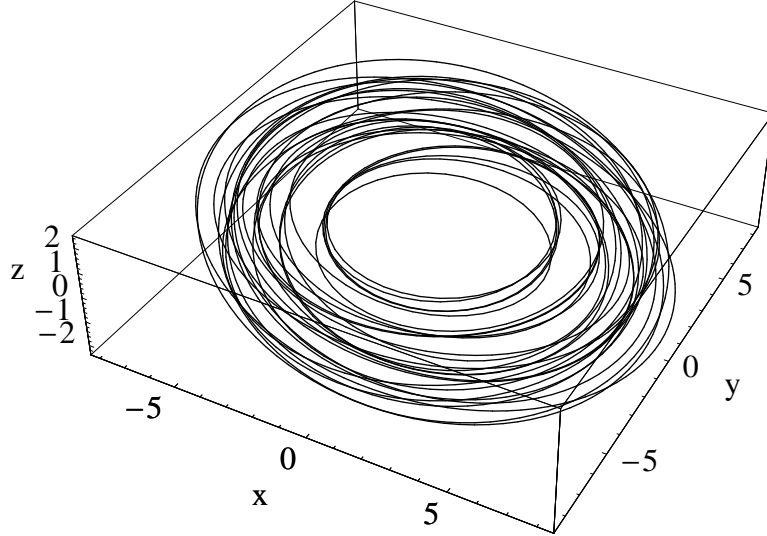


Figure 4.21: The chaotic quasi-circular orbit of two maximally spinning $10 M_{\odot}$ black holes (in the effective one body approach), corresponding to a gravitational wave frequency of $f_{\text{GW}} = 160$ Hz. The initial spins are the same as in Fig. 4.11, while the initial radius is determined by Eq. (4.29). The corresponding Lyapunov plot is shown in Fig. 4.19. Note that the quasi circularity has broken down; the radius is not even approximately constant. This qualitative change in the orbit accompanies the onset of chaos as the frequency increases (with a corresponding decrease in radius), as illustrated in Fig. 4.18.

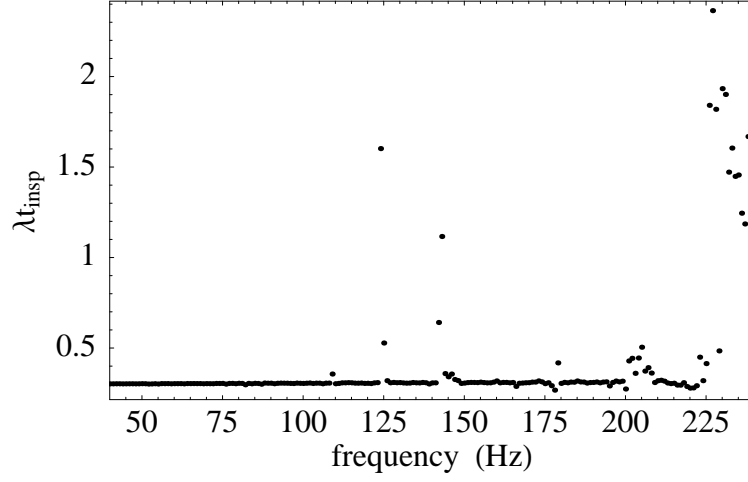


Figure 4.22: Lyapunov exponents as a function of gravitational-wave frequency for quasi-circular orbits of two maximally spinning black holes, $(20+10)M_{\odot}$ configuration. The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$, and the frequencies are chosen to correspond closely to the LIGO frequency band. There are a couple of scattered chaotic orbits at lower frequencies, but most of the chaos is concentrated at high frequencies. The initial conditions are determined by the frequency-radius relation [Eq. (4.29)], the conditions for quasi circularity, and the initial spins: $\mathbf{S}_1 = (-0.935125, 0.329567, 0.130101) m_1^2$ and $\mathbf{S}_2 = (0.039523, -0.54303, -0.838783) m_2^2$.

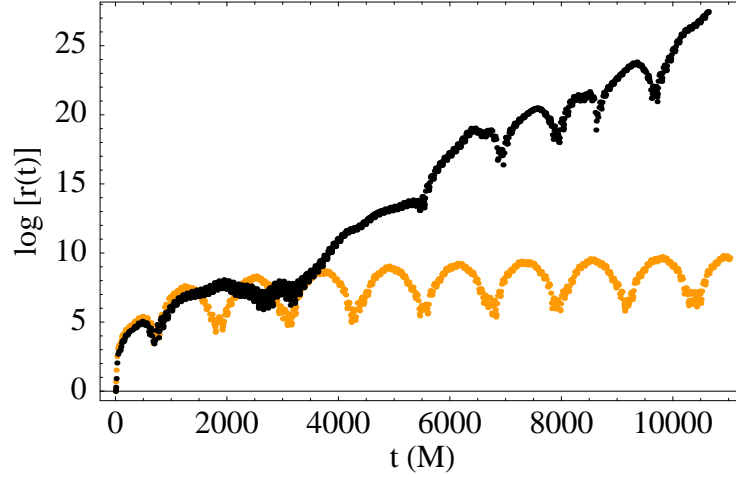


Figure 4.23: The natural logarithms of the principal ellipsoid axes vs. time for frequencies on opposite sides of the transition to chaos shown in Fig. 4.22: $f_{\text{GW}} = 224$ Hz (light) and $f_{\text{GW}} = 227$ Hz (dark). The slope of each line is the Lyapunov exponent, with $\lambda \approx 2.9 \times 10^{-4} M^{-1}$ (light, non-chaotic/consistent with zero) and $\lambda \approx 2.2 \times 10^{-3} M^{-1}$ (dark, chaotic). The initial spins are the same as in Fig. 4.22, while the initial radius is determined by Eq. (4.29). The orbits corresponding to these two frequencies appear in Figs. 4.24 and 4.25.

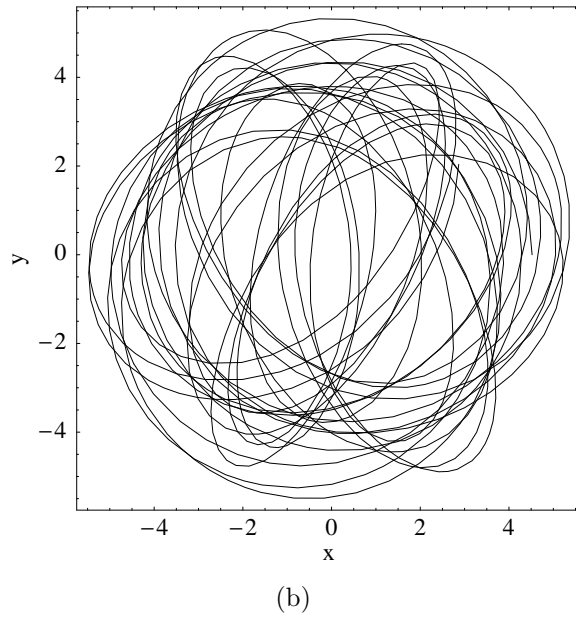
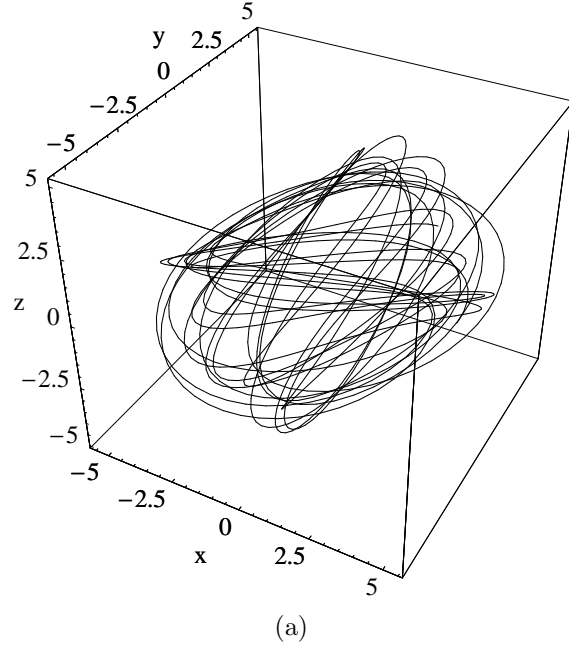
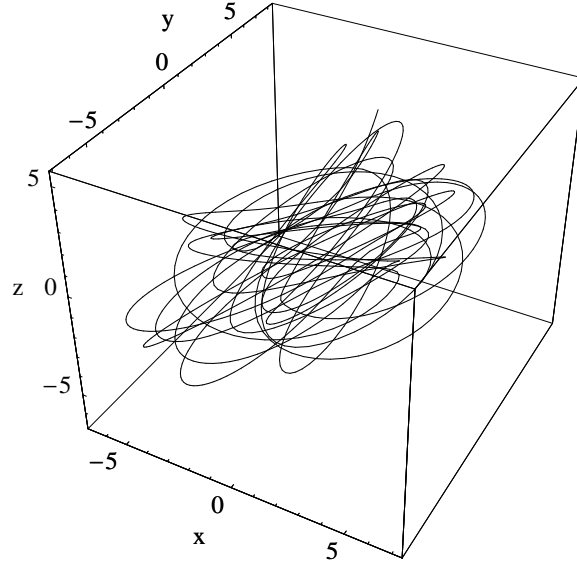
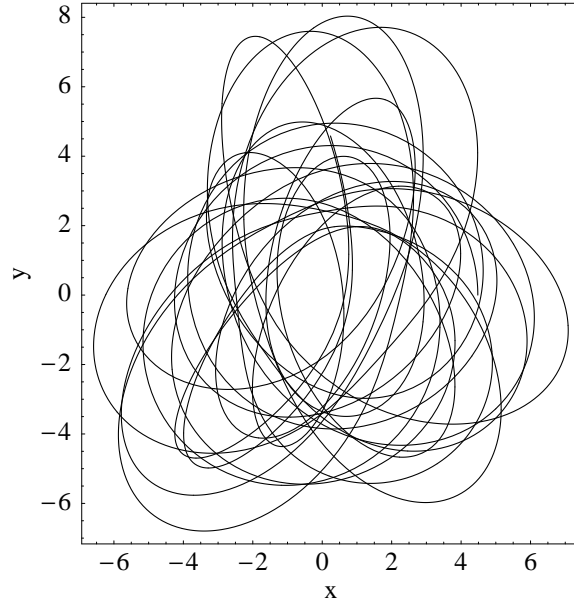


Figure 4.24: The orbit of two maximally spinning black holes, $(20 + 10)M_{\odot}$ configuration. (a) The orbit embedded in Euclidean space; (b) the projection onto the x - y plane. Lengths are measured in terms of the total mass $M = m_1 + m_2$. The frequency is 224 Hz, and the spin is the same as in Fig. 4.22. This orbit gives rise to the lower (nonchaotic) curve in Fig. 4.23. Compare with Fig. 4.25.



(a)



(b)

Figure 4.25: The orbit of two maximally spinning black holes, $(20 + 10)M_{\odot}$ configuration. (a) The orbit embedded in Euclidean space; (b) the projection onto the x - y plane. Lengths are measured in terms of the total mass $M = m_1 + m_2$. The frequency is 227 Hz, and the spin is the same as in Fig. 4.22. This orbit gives rise to the upper (chaotic) curve in Fig. 4.23. The orbit is much more erratic, with a smaller pericenter, than the corresponding nonchaotic case at slightly lower frequency (Fig. 4.24).

Table 4.4: The prevalence of chaos in post-Newtonian quasi-circular orbits at 40 Hz, for spin directions chosen randomly on a unit sphere. We calculate the fraction of orbits whose e -folding times $t_\lambda = 1/\lambda$ are less than the inspiral time t_{insp} , which is our operational definition of chaos. The final integration time is 10 times the inspiral time. We also include 95% confidence intervals for the reported fractions. The simulation data represent 500 randomly chosen initial spin directions for each configuration, with the initial radius fixed by requiring a gravitational wave frequency of 40 Hz (as determined by the Newtonian formula). In this case, the number of chaotic orbits for each configuration is zero.

Configuration	Fraction chaotic	95% confidence interval
$(20 + 5)M_\odot$	0	[0, 0.00738]
$(10 + 5)M_\odot$	0	[0, 0.00738]
$(5 + 5)M_\odot$	0	[0, 0.00738]
$(10 + 10)M_\odot$	0	[0, 0.00738]
$(20 + 20)M_\odot$	0	[0, 0.00738]
$(20 + 10)M_\odot$	0	[0, 0.00738]
$(15 + 5)M_\odot$	0	[0, 0.00738]

$S_2 = 0.53$ for the $(20 + 10)M_\odot$ case. Since the dynamics in both cases are nonchaotic when $S_2 = 0$, the chaos must be produced by the spin terms in the Hamiltonian.

In Figs. 4.28 and 4.29, we show Lyapunov plots for orbits on either side of the chaotic transition. Although the difference is not as dramatic as the frequency-induced transition (Figs. 4.19 and 4.23), there is still a qualitative change in the value of the principal Lyapunov exponent. Unlike the system in Sec. 4.5.2.1, this transition does not give rise to a qualitative change in the orbit as the spin is varied. Instead, the chaos manifests itself primarily in the time-evolution of the spins, as shown in Figs. 4.30 and 4.31 [the $(10 + 10)M_\odot$ configuration] and Figs. 4.32 and 4.33 [the $(20 + 10)M_\odot$ configuration].

4.5.2.4 Varying the PN terms

The previous results in this section included all the PN terms described in Sec. 4.2.1. Here we investigate the effect on the presence of chaos of turning off some of the terms. We focus on a chaotic orbit found in our simulations of the $(20 + 10)M_\odot$ configuration for f_{GW} , as shown in Fig. 4.14. The Lyapunov plots for a variety of PN term combinations appear in Fig. 4.34. The most important result is that the spin-spin terms are crucial to the presence of chaos; when only Newtonian, 1PN, 2PN, and spin-orbit are turned on, the system is not chaotic. The only surprising result is that the $S_1 S_1$ quadrupole term apparently exerts a stabilizing influence: when only $S_2 S_2$ is turned off, the system is nonchaotic, but if $S_1 S_1$ is then turned off as well (leaving $S_1 S_2$ as the only spin-spin term) the system returns to chaotic behavior.

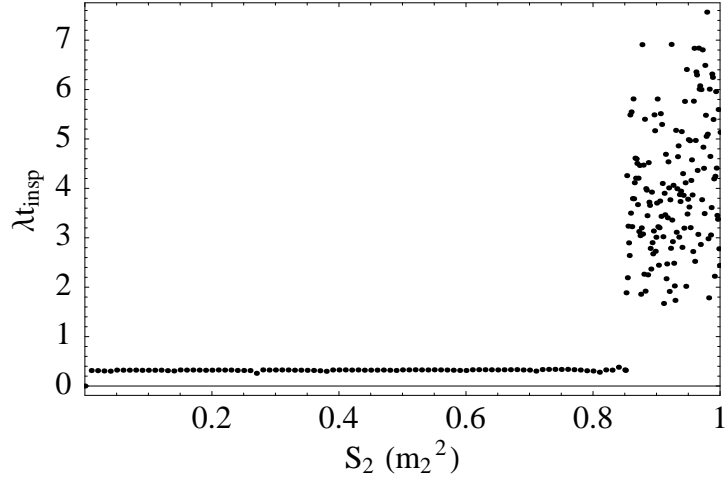


Figure 4.26: Lyapunov exponents as a function of spin for quasi-circular orbits of two $10 M_\odot$ black holes at 240 Hz. We fix the spin of one hole at the maximum value ($S_1 = m_1^2$), and also fix the spin directions (as given in Fig. 4.11), and then vary the spin S_2 of the second body. The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$. There is an abrupt transition to chaos when S_2 exceeds 0.85 (measured in units of m_2^2).

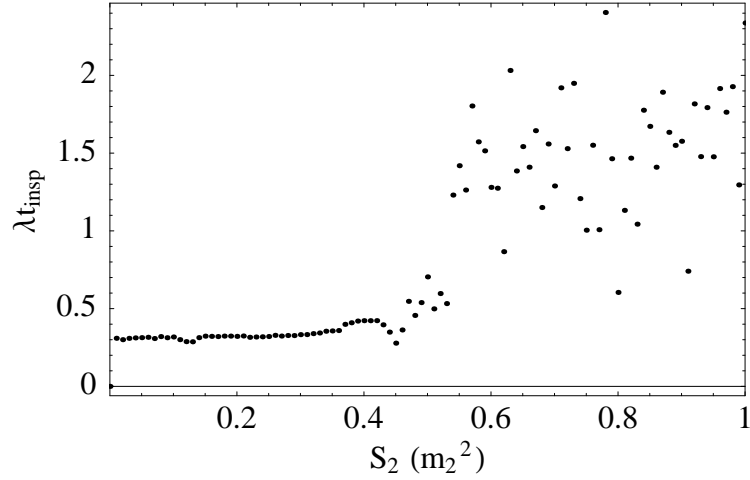


Figure 4.27: Lyapunov exponents as a function of spin for quasi-circular orbits of two spinning black holes at 240 Hz, $(20 + 10)M_\odot$ configuration. The initial spins are in the same directions as in Fig. 4.22, while the initial radius is determined by Eq. (4.29). We fix the spin of one hole at the maximum value ($S_1 = m_1^2$), and also fix the spin directions (as given in Fig. 4.22), and then vary the spin S_2 of the second body. The Lyapunov exponents are measured in terms of the inverse inspiral time $1/t_{\text{insp}}$. There is a transition to chaos when S_2 exceeds 0.53 (measured in units of m_2^2); the values of λt_{insp} for lower values of S_2 are indistinguishable from zero (i.e., they are the same size as the exponent for a nonchaotic baseline orbit).

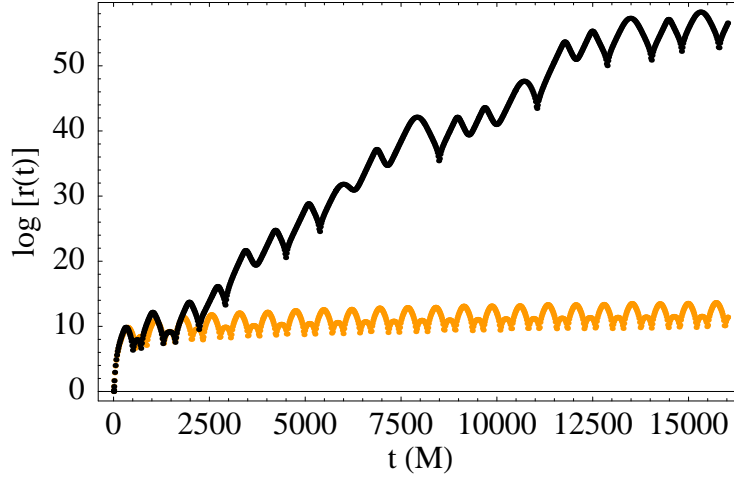


Figure 4.28: The natural logarithms of the principal ellipsoid axis vs. time for spins on opposite sides of the transition to chaos shown in Fig. 4.26: $S_2 = 0.83$ (light) and $S_2 = 0.86$ (dark). The slope of each line is the Lyapunov exponent, with $\lambda \approx 2.0 \times 10^{-4} M^{-1}$ (light, nonchaotic/consistent with zero) and $\lambda \approx 3.5 \times 10^{-3} M^{-1}$ (dark, chaotic). The orbits corresponding to these two frequencies appear in Figs. 4.30 and 4.31.

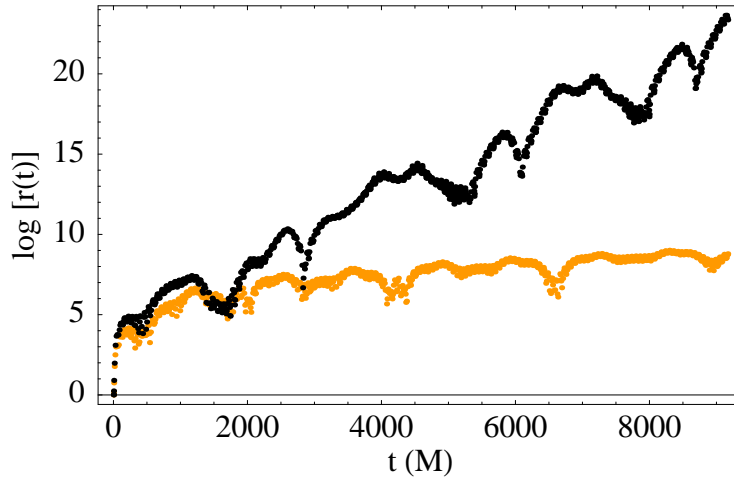
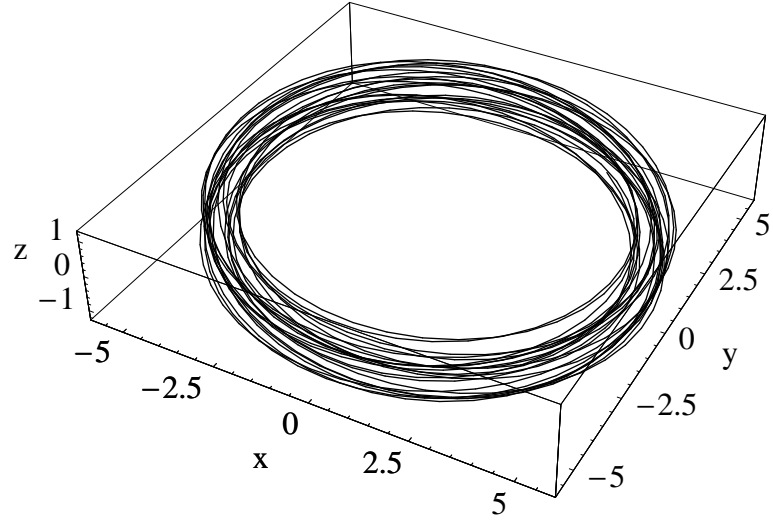
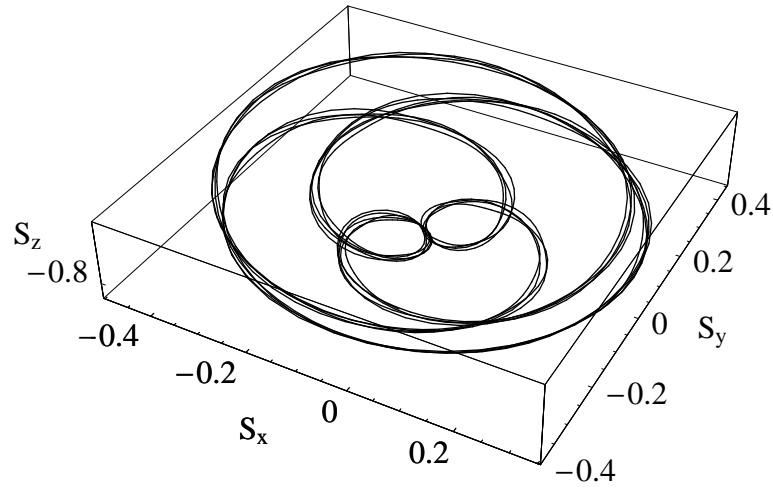


Figure 4.29: The natural logarithms of the principal ellipsoid axis vs. time for spins on opposite sides of the transition to chaos shown in Fig. 4.27: $S_2 = 0.40$ (light) and $S_2 = 0.58$ (dark). The initial spins are in the same directions as in Fig. 4.22, while the initial radius is determined by Eq. (4.29). The slope of each line is the Lyapunov exponent, with $\lambda \approx 4.6 \times 10^{-4} M^{-1}$ (light, nonchaotic/consistent with zero) and $\lambda \approx 2.0 \times 10^{-3} M^{-1}$ (dark, chaotic). The initial spins are in the same directions as in Fig. 4.22, while the initial radius is determined by Eq. (4.29). The orbits corresponding to these two frequencies appear in Figs. 4.32 and 4.33.

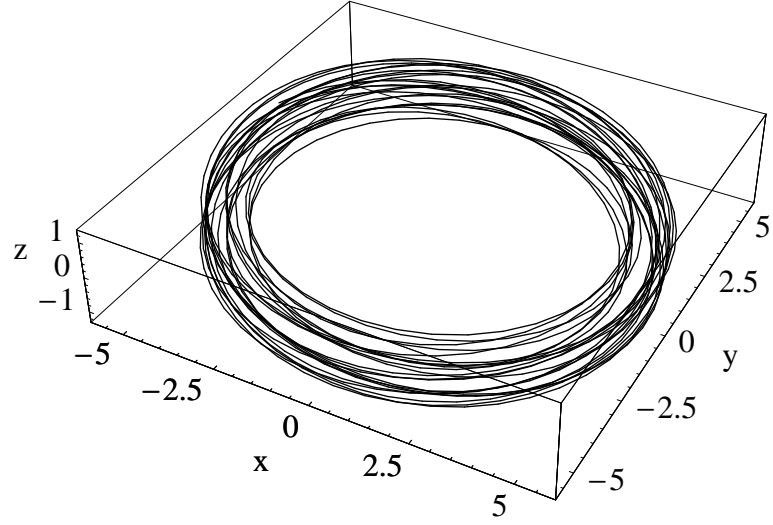


(a)

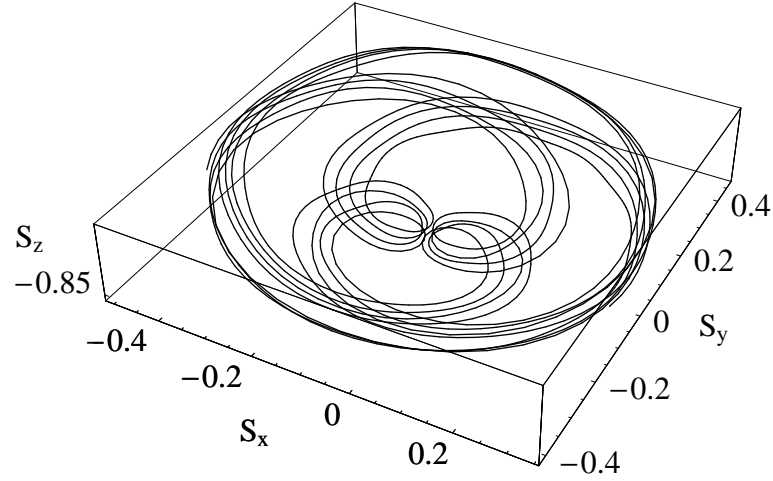


(b)

Figure 4.30: (a) The nonchaotic quasi-circular orbit of two $10 M_{\odot}$ black holes (in the effective one-body approach) at 240 Hz, with spin magnitudes of $S_1 = 1$ and $S_2 = 0.83$; (b) Cartesian “spin space” showing the time-evolution of S_x , S_y , and S_z . Compare with Fig. 4.31(b). The initial spins are in the same directions as in Fig. 4.11, while the initial radius is determined by Eq. (4.29). The corresponding Lyapunov plot is shown in Fig. 4.28.

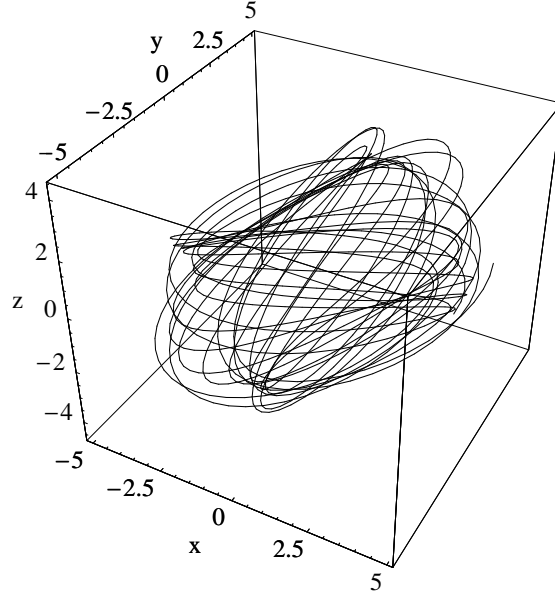


(a)

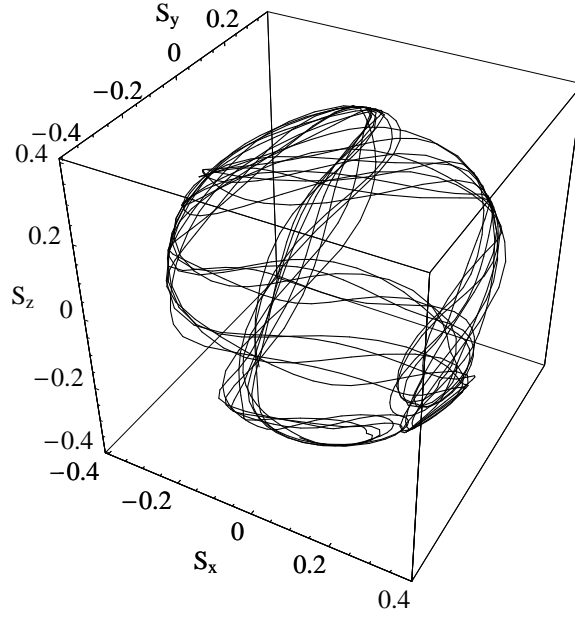


(b)

Figure 4.31: (a) The chaotic quasi-circular orbit of two $10 M_\odot$ black holes (in the effective one-body approach), with spin magnitudes of $S_1 = 1$ and $S_2 = 0.86$ at 240 Hz; (b) Cartesian “spin space” showing the time-evolution of S_x , S_y , and S_z . The initial spins are in the the same direction as in Fig. 4.11, while the initial radius is determined by Eq. (4.29). The corresponding Lyapunov plot is shown in Fig. 4.28. There appears to be no qualitative difference between the orbit (a) and the orbit in Fig. 4.30(a), but there is a qualitative change in the spin behavior. Unlike the frequency transition shown in Fig. 4.19, which gives a qualitative change in the evolution of the spatial variables, the spin transition to chaos manifests itself in the spin degrees of freedom.

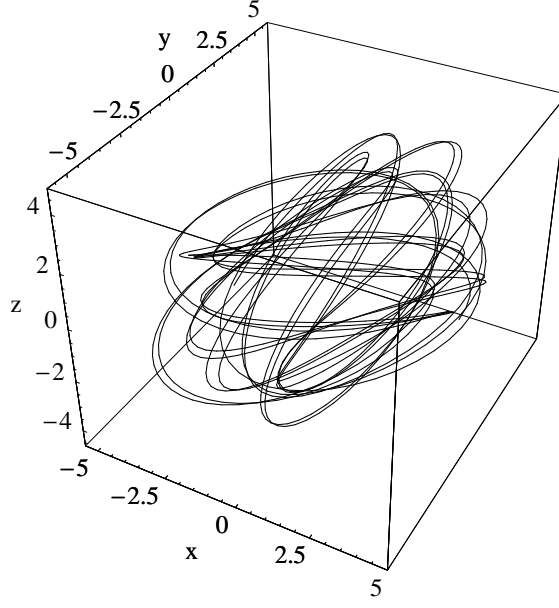


(a)

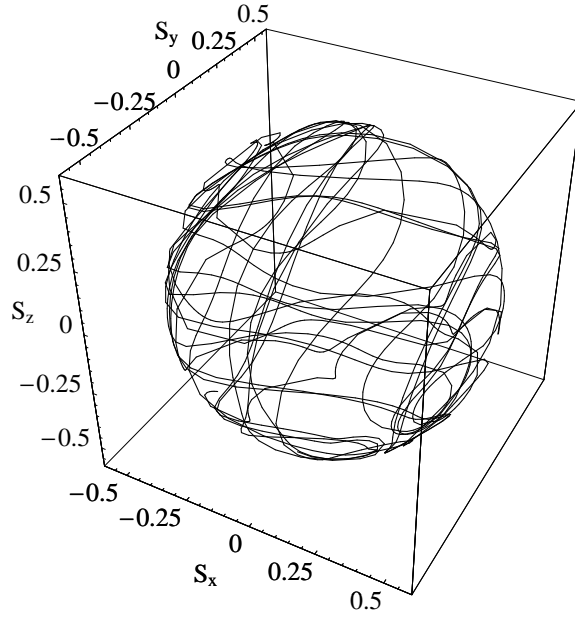


(b)

Figure 4.32: (a) The nonchaotic quasi-circular orbit of two black holes, $(20 + 10)M_\odot$ configuration (in the effective one-body approach) at 240 Hz, with spin magnitudes of $S_1 = 1$ and $S_2 = 0.40$; (b) Cartesian “spin space” showing the time-evolution of S_x , S_y , and S_z . Compare with Fig. 4.33(b). The initial spins are in the same directions as in Fig. 4.22, while the initial radius is determined by Eq. (4.29). The corresponding Lyapunov plot is shown in Fig. 4.29.



(a)



(b)

Figure 4.33: (a) The chaotic quasi-circular orbit of two black holes, $(20 + 10)M_\odot$ configuration (in the effective one-body approach), with spin magnitudes of $S_1 = 1$ and $S_2 = 0.58$ at 240 Hz; (b) Cartesian “spin space” showing the time-evolution of S_x , S_y , and S_z . The initial spins are in the same directions as in Fig. 4.22, while the initial radius is determined by Eq. (4.29). The corresponding Lyapunov plot is shown in Fig. 4.29. There is little difference between the orbit (a) and the orbit in Fig. 4.32(a), but there is a qualitative change in the spin behavior. Unlike the frequency transition shown in Fig. 4.23, which gives a qualitative change in the evolution of the spatial variables, the spin transition to chaos manifests itself primarily in the spin degrees of freedom.

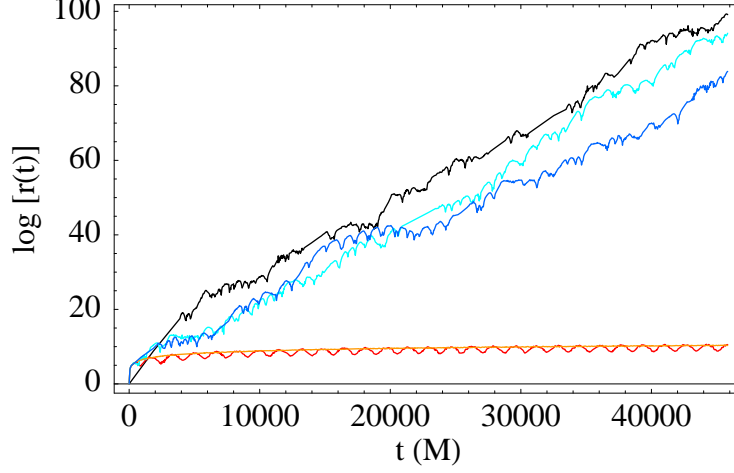


Figure 4.34: The natural logarithms of the longest ellipsoid axis r vs. t for a strongly chaotic $(20+10)M_{\odot}$ configuration, for varying post-Newtonian terms. The initial spins (both maximal) are $\mathbf{S}_1 = (-0.935125, 0.329567, 0.130101)m_1^2$ and $\mathbf{S}_2 = (0.039523, -0.54303, -0.838783)m_2^2$, while the initial radius is $r = 4.318M$, corresponding to $f_{\text{GW}} = 240$ Hz; the other initial conditions are fixed by the conditions for quasi circularity (Sec. 4.3.2). The final time is 50 times the inspiral time, and the slopes of the lines are the Lyapunov exponents. In all cases, the Newtonian, 1PN, 2PN, and spin-orbit terms are turned on, but some of the others may be turned off; from top to bottom (with colors, visible in electronic versions of this chapter, noted parenthetically): all PN terms (black); S_1S_1 and S_2S_2 turned off (cyan); S_1S_1 turned off (blue); S_2S_2 turned off (wiggly/red, nonchaotic); S_1S_1 , S_2S_2 , and S_1S_2 turned off (straight/orange, nonchaotic).

4.6 Conclusions and future work

The dynamics of binary black holes, as modeled by the post-Newtonian equations, are significantly affected by the presence of spin. In particular, the addition of spin terms to the post-Newtonian equations of motion leads to significant changes in the orbital geometry and dynamical behavior of the solutions. The effects of the spin terms are particularly clear on quasi-circular orbits, where the interaction terms quadratic in the spin cause deviations from perfectly spherical orbits. We also find that, in the extreme mass-ratio limit, the PN equations agree well with the Papapetrou equations, for both eccentric and near-circular orbits.

We find that, for quasi-circular orbits, the presence of the interaction terms quadratic in the spins leads to chaotic solutions for a variety of different parameter values, as indicated by positive Lyapunov exponents. These exponents come in $\pm\lambda$ pairs, a reflection of the Hamiltonian nature of the dynamics. We measure the strength of the chaos by comparing the e -folding timescale for chaotic behavior (the inverse of the Lyapunov exponent) with the inspiral timescale. We find especially strong chaos for high-frequency/low-radius orbits and high spins. By varying the frequency and spin parameters, we find that a smooth change of parameters leads to an abrupt transition to chaos, accompanied by qualitative changes in the solutions to the equations of motion.

Despite the unambiguous presence of chaos, chaotic solutions are rare, occurring only for a small percentage of initial conditions even at the highest frequencies and spins. Moreover, the “quasi-circular” orbits producing positive Lyapunov exponents, which are not even approximately circular due to the strong spin-spin coupling, have pericenters that are not small compared to M . In other words, the equations of motion are chaotic only for orbits where the post-Newtonian approximation (which requires $r \ll M$) is invalid. Finally, for the vast majority of the combinations of frequency and spin—even those at high frequencies with maximal spins—chaos is entirely absent.

There are several natural directions for future work. The addition of the 3PN term to the Hamiltonian should help indicate whether the chaos discovered in this study is a result of a breakdown in the equations themselves. It would also be valuable to generalize the survey of quasi-circular orbits to eccentric orbits. Finally, a study of neutron star binaries that includes the spin-induced quadrupole terms (calculated using a parameterized equation of state) would be useful for understanding these important systems.

Acknowledgements

Thanks to Alessandra Buonanno for her encouragement and for supplying the background material that made this project possible. The material in Sec. 4.2 and much of Sec. 4.3 is based on notes that she provided. I also thank Yanbei Chen for several helpful conversations. Finally, thanks to Sterl Phinney for his support and feedback. This work was supported in part by NASA grant NAG5-10707.

Bibliography

- [1] S. Suzuki and K. Maeda, Phys. Rev. D **55**, 4848 (1997).
- [2] J. Levin, Phys. Rev. Lett. **84**, 3515 (2000).
- [3] J. D. Schnittman and F. A. Rasio, Phys. Rev. Lett. **87**, 121101 (2001).
- [4] N. J. Cornish and J. Levin, [gr-qc/0207016](#).
- [5] M. D. Hartl, Phys. Rev. D **67**, 024005 (2003), [gr-qc/0210042](#).
- [6] B. J. Owen, Phys. Rev. D **53**, 6749 (1996).
- [7] E. E. Flanagan and S. A. Hughes, Phys. Rev. D **57**, 4566 (1998).
- [8] A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **67**, 024016 (2003).
- [9] N. J. Cornish, Phys. Rev. D **64**, 084011 (2001).
- [10] T. Damour, Phys. Rev. D **64**, 124013 (2001).
- [11] T. Damour and G. Schäfer, Nuov. Cimento **101**, 127 (1988).
- [12] T. Damour, P. Jaranowski, and G. Schäfer, Phys. Rev. D **62**, 084011 (2000).
- [13] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
- [14] F. D. Ryan, Phys. Rev. D **52**, 3159 (1995).
- [15] C. M. Will, Living Rev. Relativity **4** (30 Apr 2003),
<http://www.livingreviews.org/Articles/Volume4/2001-4will>.
- [16] W. G. Dixon, Proc. R. Soc. London **A314**, 499 (1970).
- [17] I. B. Khriplovich and A. A. Pomeransky (1996), [gr-qc/9602004](#).
- [18] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
- [19] J. Levin, Phys. Rev. D **67**, 044013 (2003).

- [20] N. J. Cornish and J. Levin, Phys. Rev. Lett. **89**, 179001 (2002).
- [21] N. Cornish and J. Levin, Phys. Rev. Lett. **78**, 998 (1997).
- [22] P. C. Peters, Phys. Rev. **136**, 1224 (1964).
- [23] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Phys. Rev. D **49**, 6274 (1994).

Chapter 5

Lyapunov exponents in constrained and unconstrained ODEs

Abstract

We discuss several numerical methods for calculating Lyapunov exponents (a quantitative measure of chaos) in systems of ordinary differential equations. We pay particular attention to constrained systems, and we introduce a variety of techniques to address the complications introduced by constraints. For all cases considered, we develop both deviation vector methods, which follow the time-evolution of the difference between two nearby trajectories, and Jacobian methods, which use the Jacobian matrix to determine the true local behavior of the system. We also assess the merits of the various methods, and discuss assorted subtleties and potential sources of error.

5.1 Introduction

Chaos exists in a wide variety of nonlinear mathematical and physical systems, and ordinary differential equations are no exception. Since the original discovery by Edward Lorenz of deterministic chaos in a toy atmosphere model (consisting of twelve differential equations) [1], a seemingly endless variety of ODEs exhibiting extreme sensitivity to initial conditions has emerged. Many tools, both qualitative and quantitative, have been developed to investigate this chaotic behavior. Perhaps the most important quantitative measure of chaos is the method of Lyapunov exponents, which indicate the average rate of separation for nearby trajectories. (See [2–7] for some recent investigations into measures of chaos and their applications.) The present chapter is concerned with general methods for calculating these exponents in arbitrary systems of ODEs. We first review the techniques for calculating Lyapunov exponents in unconstrained systems [8, 9] (where each coordinate represents a true degree of freedom), and then introduce several new methods for calculating Lyapunov exponents in constrained systems (where there are more coordinates than there are degrees of freedom).

A defining characteristic of a chaotic dynamical system is *sensitive dependence on initial conditions*, and the Lyapunov exponents are a way of quantifying this sensitivity. In a system of ordinary differential equations, this sensitive dependence corresponds to an exponential separation of nearby phase-space trajectories: if two initial conditions are initially separated by a distance ϵ_0 , the total separation grows (on average) according to

$$\epsilon(t) = \epsilon_0 e^{\lambda t}, \quad (5.1)$$

where λ is a positive constant (with units of inverse time) called the *Lyapunov exponent*. Two important caveats to Eq. (5.1) are necessary. First, this prescription yields only the *largest* Lyapunov exponent, but a dynamical system with n degrees of freedom has in general n such exponents. Second, Eq. (5.1) does not constitute a rigorous definition, since it defines a true Lyapunov exponent only if ϵ is “infinitesimal.” A more precise definition of Lyapunov exponents involves the true local behavior of the dynamical system, i.e., the derivative or its higher-dimensional generalization.

We can go beyond Eq. (5.1) to determine (at least in principle) all n Lyapunov exponents by considering not just one nearby initial condition, but rather a ball of initial conditions with radius ϵ_0 . As discussed in Sec. 5.2, this ball evolves into an n -dimensional ellipsoid under the time-evolution of the flow, and the lengths of this ellipsoid’s principal axes determine the Lyapunov exponents. We will see that there are many advantages to this ellipsoid view, both conceptual and computational.

We discuss in Secs. 5.2 and 5.3 several techniques for calculating Lyapunov exponents in ODEs, and compare the relative merits of the various methods. We take special care to explain methods for the calculation of all n Lyapunov exponents. Our principal examples are two well-studied and simple systems: the Lorenz equations (Sec. 5.2.4.1) and the forced damped pendulum (Sec. 5.2.4.2). The techniques and code were developed and tested on the much more complex problem of spinning bodies orbiting rotating (Kerr) black holes, as discussed briefly in Sec. 5.3.4 and at length in Chapters 2 and 3.

Our two model systems are unconstrained, so that each variable represents a true degree of freedom. As we see in Sec. 5.3, following the evolution of a phase-space ellipsoid—and hence calculating the Lyapunov exponents—becomes problematic when the system is constrained. Such systems are common in physics, with constraints arising for both mathematical and physical reasons. For example, instead of using the angle θ to describe the position of a pendulum, we may find it mathematically convenient to integrate the equations of motion in Cartesian coordinates (x, y) , with a constraint on the value of $x^2 + y^2$. Another example is a spinning astronomical body, whose spin is typically described by the components of its spin vector $\mathbf{S} = (S_x, S_y, S_z)$. On physical grounds, we might wish to fix the magnitude $\|\mathbf{S}\| = S = \sqrt{S_x^2 + S_y^2 + S_z^2}$, so that only two of the three spin components represent true degrees of freedom.

We describe in Sec. 5.3 three methods for finding Lyapunov exponents in constrained systems. Our principal example of a constrained system is the forced damped pendulum described in Cartesian coordinates, a system chosen both for its conceptual simplicity and to facilitate comparison with the same system without constraints. We also show the application of these techniques to the dynamics of spinning compact objects in general relativity. It was the investigation of these constrained systems in Chapter 2 that led to the development of the key ideas described in this chapter.

We have developed a general-purpose implementation of the principal algorithms in this chapter in C++, which is available for download [10]. The user must specify the system of equations (and a Jacobian matrix if necessary), as well as a few other parameters, but the main procedures are not tied to any particular system. Most of the results in this chapter were calculated using this implementation.

We use boldface to indicate Euclidean vectors, and the symbol \log signifies the natural logarithm \log_e in all cases. We refer to the principal semiaxes of an n -dimensional ellipsoid as “axes” or “principal axes” for brevity.

5.2 Lyapunov exponents in unconstrained flows

There are two primary approaches to calculating Lyapunov exponents in systems of ordinary differential equations. The first method involves the integration of two trajectories initially separated by a small deviation vector; we obtain a measure of the divergence rate by keeping track of the length of this deviation vector. We refer to this as the *deviation vector method*. The second method uses a rigorous linearization of the equations of motion (the Jacobian matrix) in order to capture the true local behavior of the dynamical system. We call this the *Jacobian method*. Though computationally slower, the Jacobian method is more rigorous, and also opens the possibility of calculating more than just the principal exponent. In this section we discuss these two methods, and several variations on each theme, in the context of unconstrained dynamical systems.

When discussing Lyapunov exponents in ordinary differential equations, it is valuable to have both a general abstract system and a specific concrete example in mind. Abstractly, we write the coordinates of the system as a single n -dimensional vector \mathbf{y} that lives in the n -dimensional phase space, and we write the equations of motion as a system of first-order differential equations:

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}). \quad (5.2)$$

We will refer to a solution to Eq. (5.2) as a *flow*. As a specific example, consider the Lorenz system

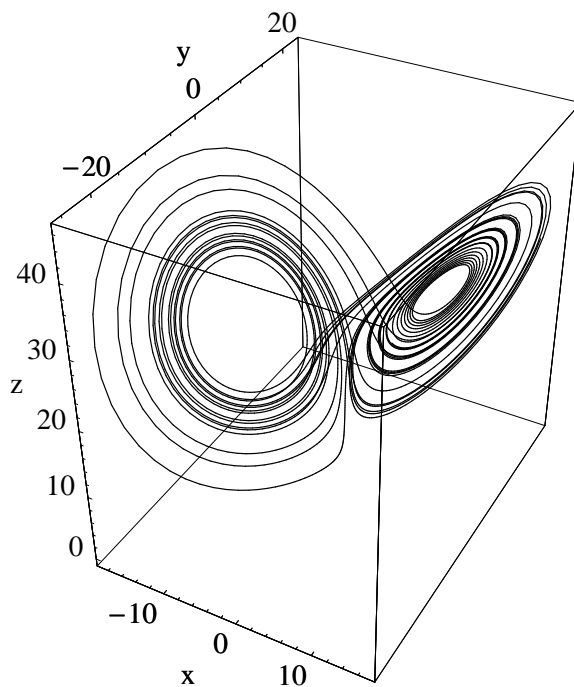


Figure 5.1: The Lorenz attractor. All initial conditions except the origin (which is an unstable equilibrium) are attracted to the figure shown.

of equations:

$$\begin{aligned}
 \dot{x} &= -\sigma x + \sigma y \\
 \dot{y} &= -xz + rx - y \\
 \dot{z} &= xy - bz,
 \end{aligned} \tag{5.3}$$

where σ , r , and b are constants. In the notation of Eq. (5.2), we then have $\mathbf{y} = (x, y, z)$ and $\mathbf{f}(\mathbf{y}) = (-\sigma x + \sigma y, -xz + rx - y, xy - bz)$. The Lorenz equations exhibit chaos for a wide variety of parameter values; in this chapter, for simplicity we consider only one such set: $\sigma = 10$, $b = 8/3$, and $r = 28$. For these parameter values, all initial conditions except the origin asymptote to the elegant Lorenz attractor (Fig. 5.1).

5.2.1 The deviation vector method

The most straightforward method for calculating the largest Lyapunov exponent is to consider an initial point $\mathbf{y}_0^{(1)} = \mathbf{y}_0$ and a nearby point $\mathbf{y}_0^{(2)} = \mathbf{y}_0 + \delta\mathbf{y}_0$, and then evolve both points forward, keeping track of the difference $\delta\mathbf{y} \equiv \mathbf{y}^{(2)} - \mathbf{y}^{(1)}$. If the motion is chaotic, then exponential separation

implies that

$$\|\delta\mathbf{y}\| = e^{\lambda_{\max} t} \|\delta\mathbf{y}_0\|, \quad (5.4)$$

so that the largest exponent is

$$\lambda_{\max} = \frac{\log[r_e(t)]}{t}, \quad (5.5)$$

where we write

$$r_e = \|\delta\mathbf{y}\|/\|\delta\mathbf{y}_0\|, \quad (5.6)$$

with a subscript e that anticipates the ellipsoid axis discussed in Sec. 5.2.2.3. Here $\|\cdot\|$ denotes the Euclidean norm (though in principle any positive-definite norm will do [11]). It is convenient to display the results of this process graphically by plotting $\log[r_e(t)]$ vs. t , which we refer to as a *Lyapunov plot*; since Eq. (5.5) is equivalent to $\log[r_e(t)] = \lambda_{\max} t$, such plots should be approximately linear, with slope equal to the principal Lyapunov exponent. (In practice, to extract the slope we perform a least-squares fit to the simulation data, which is less sensitive to fluctuations in the value of $\log[r_e(t)]$ than the ratio $\log[r_e(t_f)]/t_f$ at the final time.) We refer to this technique as the *(unrescaled) deviation vector method*.

It is important to note that, because of the problem of *saturation*, Eq. (5.5) does not define a true Lyapunov exponent. In a chaotic system, any deviation $\delta\mathbf{y}_0$, no matter how small, will eventually *saturate*, i.e., it will grow so large that it no longer represents the *local* behavior of the dynamical system. Moreover, chaotic systems are bounded by definition [in order to eliminate trivial exponential separation of the form $x(t) = x_0 e^{\lambda t}$], so there is some bound B on the distance between any two trajectories. As a result, in the infinite time limit Eq. (5.5) gives

$$\lambda_{\max} = \lim_{t \rightarrow \infty} \frac{\log \|\delta\mathbf{y}\|/\|\delta\mathbf{y}_0\|}{t} \leq \lim_{t \rightarrow \infty} \frac{\log B/\|\delta\mathbf{y}_0\|}{t} = 0. \quad (5.7)$$

In the naïve unrescaled deviation vector method, the calculated exponent is always zero because of saturation.

One solution to the saturation problem is to *rescale* the deviation once it grows too large. For example, suppose that we set $\|\delta\mathbf{y}_0\| = \epsilon$ for some small ϵ (say 10^{-8}), and then allow the deviation to grow by at most a factor of f . Then, whenever $\|\delta\mathbf{y}\| \geq f \|\delta\mathbf{y}_0\|$, we rescale the deviation back to a size ϵ and record the length $R_i = \|\delta\mathbf{y}\|/\|\delta\mathbf{y}_0\|$ of the expanded vector. If we perform N such rescalings in the course of a calculation, the total expansion of the initial vector is then

$$r_e = \frac{\|\delta\mathbf{y}_f\|}{\|\delta\mathbf{y}_0\|} \prod_{i=1}^N R_i, \quad (5.8)$$

where $\delta\mathbf{y}_f$ is the final size of the (rescaled) separation vector. Applying Eq. (5.5), we see that the

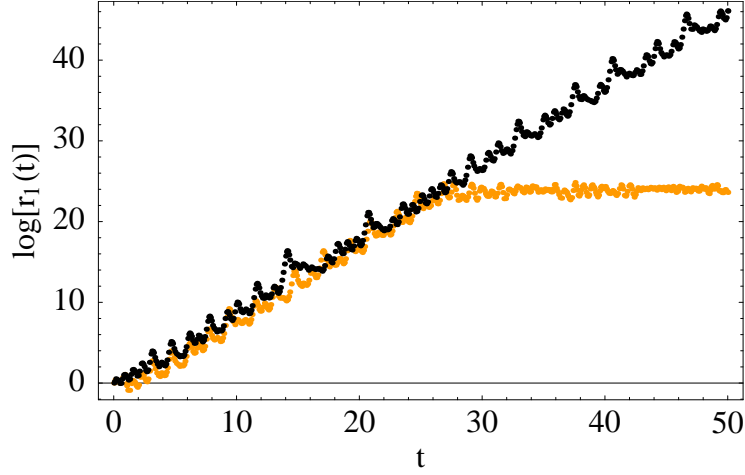


Figure 5.2: Comparison of the unrescaled (light) and rescaled (dark) deviation vector methods for calculating the principal Lyapunov exponent of the Lorenz system [Eq. (5.3)]. The slope of the rescaled line is the Lyapunov exponent ($\lambda_{\max} = 0.905 \pm 0.003$; see Sec. 5.2.4.1). The initial deviation is $\|\delta \mathbf{y}_0\| = 10^{-8}$, and rescaling occurs (for the rescaled method) if $\|\delta \mathbf{y}\| \geq 10^{-2}$. Note the saturation of the unrescaled approach once the deviation has grown too large.

approximate Lyapunov exponent satisfies

$$\lambda_{\max} = \frac{1}{t} \left[\log \left(\frac{\|\delta \mathbf{y}_f\|}{\|\delta \mathbf{y}_0\|} \right) + \sum_{i=1}^N \log R_i \right]. \quad (5.9)$$

We refer to this as the *(rescaled) deviation vector method*.

The rescaled deviation vector method is not particularly robust compared to the rigorous method described below (Sec. 5.2.2), and there are significant complications when applying it to constrained systems, but if implemented with care it provides a fast and accurate estimate for the largest Lyapunov exponent. Fig. 5.2 shows both the rescaled and unrescaled deviation vector methods applied to the Lorenz system [Eq. (5.3)]. Note in particular the saturation of the unrescaled approach. We discuss the limitations of the rescaled method further in Sec. 5.4.

5.2.2 The Jacobian method

Although the deviation vector method suffices for practical calculation in many cases, in essence it amounts to taking a numerical derivative. For a one-dimensional function of one variable, we can approximate the derivative at $x = x_0$ using

$$f'(x_0) \approx \frac{f(x_0 + \epsilon) - f(x_0)}{\epsilon}, \quad (5.10)$$

for some $\epsilon \ll 1$, but this prescription is notoriously inaccurate as a numerical calculation [12]. Of course, it is better (if possible) to calculate the analytical derivative $f'(x)$ and evaluate it at x_0 . The higher-dimensional generalization of this is the Jacobian matrix, which describes the local (linear) behavior of a higher-dimensional function. In the context of a dynamical system, this means that we can find the time-evolution of a small deviation $\delta\mathbf{y}$ using the rigorous linearization of the equations of motion:

$$\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) = \mathbf{Df} \cdot \delta\mathbf{y} + O(\|\delta\mathbf{y}\|^2), \quad (5.11)$$

where

$$(\mathbf{Df})_{ij} = \frac{\partial f_i}{\partial x^j} \quad (5.12)$$

is the Jacobian matrix evaluated along the flow. For example, for the Lorenz system [Eq. (5.3)] we have

$$\mathbf{Df} = \begin{pmatrix} -\sigma & \sigma & 0 \\ r - z(t) & -1 & -x(t) \\ y(t) & x(t) & -b \end{pmatrix}, \quad (5.13)$$

where we write the coordinates as functions of time to emphasize that Eq. (5.13) is different at each time t .

5.2.2.1 Jacobian diagnostic

One note about Jacobian matrices is worth mentioning: practical experience has shown that errors occasionally creep into the calculations leading to the Jacobian matrix, especially if the equations of motion are complicated. It is therefore worthwhile to note that Eq. (5.11) provides an invaluable diagnostic: calculate the quantity

$$\Delta = \mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) - \mathbf{Df} \cdot \delta\mathbf{y} \quad (5.14)$$

for varying values of $\|\delta\mathbf{y}\|$; if Δ does not generally scale as $\|\delta\mathbf{y}\|^2$, then something is amiss. (The routines in [10] include this important Jacobian diagnostic function.)

5.2.2.2 The principal exponent

The main value of Eq. (5.11) in the context of a dynamical system is its combination with Eq. (5.2) to yield an equation of motion for the deviation $\delta\mathbf{y}$:

$$\mathbf{f}(\mathbf{y} + \delta\mathbf{y}) = \frac{d}{dt}(\mathbf{y} + \delta\mathbf{y}) = \mathbf{f}(\mathbf{y}) + \frac{d(\delta\mathbf{y})}{dt}, \quad (5.15)$$

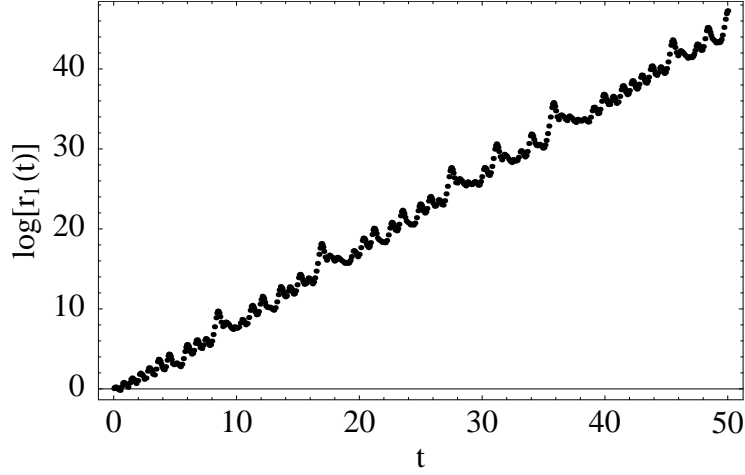


Figure 5.3: The natural logarithm of the tangent vector length $r_1 \equiv \|\xi(t)\|$ vs. t for the Lorenz system. The slope of the rescaled line is the system's largest Lyapunov exponent ($\lambda_{\max} \approx 0.905$). The figure and exponent are virtually identical to the rescaled deviation method show in Fig. 5.2.

so that (discarding terms higher than linear order) Eq. (5.11) gives

$$\frac{d(\delta \mathbf{y})}{dt} = \mathbf{Df} \cdot \delta \mathbf{y}. \quad (5.16)$$

This equation is only approximately true for finite (that is, non-infinitesimal) deviations, but we can take the infinitesimal limit by identifying the deviation $\delta \mathbf{y}$ with an element ξ in the tangent space at \mathbf{y} . This leads to an exact equation for ξ :

$$\frac{d\xi}{dt} = \mathbf{Df} \cdot \xi. \quad (5.17)$$

The initial value of ξ is arbitrary, but it is convenient to require that $\|\xi_0\| = 1$, so that the factor by which ξ has grown at some later time t is simply $\|\xi(t)\|$.

The core of the *Jacobian method* for the principal Lyapunov exponent is to solve Eqs. (5.2) and (5.17) as a coupled set of differential equations. As in Sec. 5.2.1, for chaotic systems the length of the deviation vector will grow exponentially, so that

$$\|\xi(t)\| \approx e^{\lambda_{\max} t}, \quad (5.18)$$

which implies that

$$\lambda_{\max} = \frac{\log \|\xi(t)\|}{t}. \quad (5.19)$$

For sufficiently large values of t , Eq. (5.19) provides an approximation for the largest Lyapunov exponent. It is essential to understand that there is no restriction on the length of the tangent

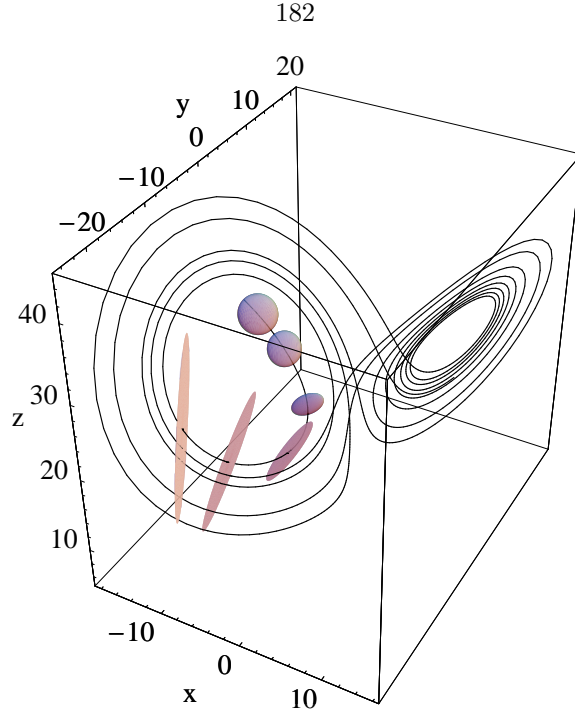


Figure 5.4: The Lorenz system with an evolving ellipsoid. The ellipsoid is calculated exactly in the tangent space (for a total time $t = 0.4$) and is superposed on the phase space for the purposes of visualization. There is one expanding axis ($\sim e^{0.905t}$) and one contracting axis ($\sim e^{-14.57t}$); the third axis has a fixed unit length (Sec. 5.2.4.1).

vector ξ : the Jacobian method *does not saturate*. The only limitation on the size of ξ in practice is the maximum representable floating point number on the computer.

5.2.2.3 Ellipsoids and multiple exponents

Although following the time-evolution of a tangent vector ξ in place of a finite deviation $\delta\mathbf{y}$ solves the problem of saturation, it still only allows us to determine the principal exponent λ_{\max} . For a system with n degrees of freedom, this leaves $n - 1$ exponents undetermined. In order to calculate all n exponents, we must introduce n tangent vectors. (We discuss the value of knowing all n exponents in Sec. 5.2.3 below.) Since n (linearly independent) vectors span an n -dimensional ellipsoid, this leads to a visualization of the Lyapunov exponents in terms of the evolution of a tangent space ellipsoid (Fig. 5.4). Fig. 5.5 shows the corresponding Lyapunov plot.

The general method is to introduce a linearly independent set of vectors $\{\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(n)}\}$. It is convenient to begin the integration with vectors that form the orthogonal axes of a unit ball, so that the vectors $\{\xi_0^{(1)}, \xi_0^{(2)}, \dots, \xi_0^{(n)}\}$ are orthonormal. Each of these tangent vectors satisfies its own version of Eq. (5.17):

$$\frac{d\xi^{(n)}}{dt} = \mathbf{Df} \cdot \xi^{(n)}. \quad (5.20)$$

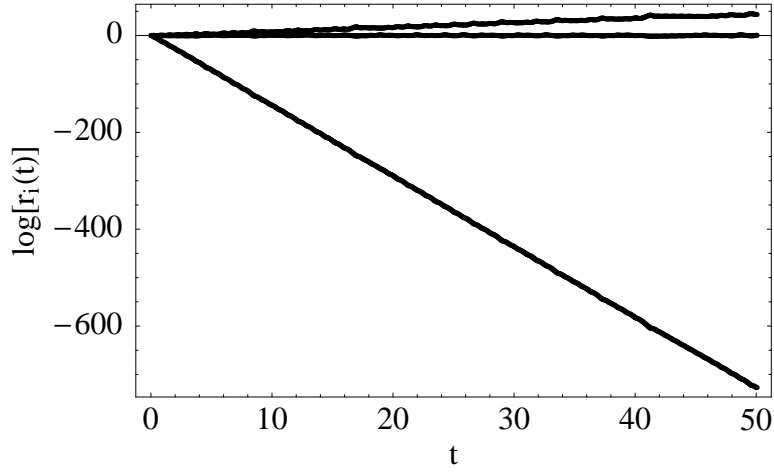


Figure 5.5: The natural logarithms of all three of the ellipsoid axes r_i vs. t for the Lorenz system, calculated using the Jacobian method (Sec. 5.2.2). The slopes are the Lyapunov exponents. The three lines correspond to the exponents $\lambda_1 \approx 0.905$, $\lambda_2 \approx 0.0$, and $\lambda_3 \approx -14.57$ (Sec. 5.2.4.1). These values agree with the calculations in [9].

If we combine the n tangent vectors to form the columns of a matrix \mathbf{U} , then Eq. (5.20) implies that

$$\frac{d\mathbf{U}}{dt} = \mathbf{Df} \cdot \mathbf{U}. \quad (5.21)$$

This equation, combined with Eq. (5.2), describes the evolution of a unit ball into an n -dimensional ellipsoid.

The value of the tangent space ellipsoid is this: if r_i is the i th principal ellipsoid axis [and $r_i(0) = 1$], then

$$r_i(t) = e^{\lambda_i t}, \quad (5.22)$$

where λ_i is the i th Lyapunov exponent. That is, the ellipsoid's axes grow (or shrink) exponentially, and if $\lambda_i > 0$ for any i then the system is chaotic [11]. [Recall that we refer to the semiaxes as “axes” for brevity (Sec. 5.1).] Turning Eq. (5.22) around, we can find the i th Lyapunov exponent by finding the average stretching (or shrinking) per unit time of the i th principal ellipsoid axis:

$$\lambda_i \approx \frac{\log[r_i(t)]}{t}. \quad (5.23)$$

In practice, a more robust prescription is to record $\log[r_i(t)]$ as a function of t and perform a least-squares fit to the pairs $(t_j, \log[r_i(t_j)])$ to find the slope λ_i .

Though Eq. (5.23) provides an estimate for the i th Lyapunov exponent, it requires us to find the n principal axes of the final ellipsoid. While it is true that the columns of the final matrix \mathbf{U}_f necessarily span an ellipsoid, but they are not in general orthogonal; in particular, the final tangent

vectors do not necessarily coincide with the ellipsoid's principal axes. A first step in extracting these axes is to note an important theorem in linear algebra (see [9] for a proof):

Theorem 1 *Let A be an $n \times n$ real matrix consisting of n linearly independent column vectors $\{\mathbf{v}_i\}_{i=1}^n$, and let $\{s_i^2\}_{i=1}^n$ be the eigenvalues and $\{\mathbf{u}_i\}_{i=1}^n$ the normalized eigenvectors of $A^T A$ (where A^T is the transpose of A). Then $\{\mathbf{v}_i\}_{i=1}^n$ lie on an n -dimensional ellipsoid whose principal axes are $\{s_i \mathbf{u}_i\}_{i=1}^n$.*

In other words, finding the principal axes of the ellipsoid represented by a matrix A is equivalent to finding the eigensystem of $A^T A$. (We note that the ellipsoid is unique: any other matrix B whose columns $\{\mathbf{w}_i\}_{i=1}^n$ lie on the same ellipsoid as $\{\mathbf{v}_i\}_{i=1}^n$ must necessarily give the same principal axes.)

In principle, we are done: simply evolve \mathbf{U} for a long time, and find the eigenvalues of $\mathbf{U}^T \mathbf{U}$. In practice, this fails miserably; every (generic) initial vector $\boldsymbol{\xi}_0^{(i)}$ has some component along the direction of greatest stretching, so *all* initial tangent space vectors eventually point approximately along the longest principal axis. As a result, all axes but the longest one are lost due to finite floating point precision.

The solution is to find new orthogonal axes as the system evolves. In other words, we can let the system evolve for some time T , stop to calculate the principal axes of the evolving ellipsoid, and then continue the integration. The method we advocate is the Gram-Schmidt orthogonalization procedure, which results in an orthogonal set of vectors spanning the same volume as the original ellipsoid, and with directions that converge to the true ellipsoid axes. This approach, originally described in [13], is a common textbook approach [8, 9], and was used successfully in Chapter 2. Numerically, the Gram-Schmidt algorithm is subject to considerable roundoff error [12], and is usually considered a poor choice for orthogonalizing vectors, but in the context of dynamics its performance has proven to be astonishingly robust. (See Sec. 5.4 for further discussion.)

We review briefly the Gram-Schmidt construction, and then indicate its use in calculating Lyapunov exponents. Given n linearly-independent vectors $\{\mathbf{u}_i\}$, the Gram-Schmidt procedure constructs n orthogonal vectors $\{\mathbf{v}_i\}$ that span the same space, given by

$$\mathbf{v}_i = \mathbf{u}_i - \sum_{j=1}^{i-1} \frac{\mathbf{u}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_j\|^2} \mathbf{v}_j. \quad (5.24)$$

To construct the i th orthogonal vector, we take the i th vector from the original set and subtract off its projections onto the previous $i - 1$ vectors produced by the procedure. The use of Gram-Schmidt in dynamics comes from observing that the resulting vectors approximate the axes of the tangent space ellipsoid. After the first time T , all of the vectors point mostly along the principal expanding direction. We may therefore pick any one as the first vector in the Gram-Schmidt algorithm, so choose $\boldsymbol{\xi}_1 \equiv \mathbf{u}_1$ without loss of generality. If we let \mathbf{e}_i denote unit vectors along the principal axes

and let r_i be the lengths of those axes, the dynamics of the system guarantees that the first vector \mathbf{u}_1 satisfies

$$\mathbf{u}_1 = r_1 \mathbf{e}_1 + r_2 \mathbf{e}_2 + \cdots \approx r_1 \mathbf{e}_1 \equiv \mathbf{v}_1$$

since \mathbf{e}_1 is the direction of fastest stretching. The second vector \mathbf{v}_2 given by Gram-Schmidt is then

$$\mathbf{v}_2 = \mathbf{u}_1 - \frac{\mathbf{u}_1 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 \approx \mathbf{u}_1 - r_1 \mathbf{e}_1 = r_2 \mathbf{e}_2,$$

with an error of order r_2/r_1 . The procedure proceeds iteratively, with each successive Gram-Schmidt step (approximately) subtracting off the contribution due to the previous axis direction. In principle, the system should be allowed to expand to a point where $r_2 \ll r_1$, but (amazingly) in practice the Gram-Schmidt procedure converges to accurate ellipsoid axes even when the system is orthogonalized *and even normalized* on timescales short compared to the Lyapunov stretching timescale. As a result, the procedure below can be abused rather badly and still give accurate results (Sec. 5.4).

5.2.2.4 The algorithm in detail

We summarize here the method used to calculate all the Lyapunov exponents of an unconstrained dynamical system $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$ with n degrees of freedom:

1. Construct an orthonormal matrix \mathbf{U}_0 whose columns (the initial tangent vectors) span a unit ball, and then integrate

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}) \tag{5.25}$$

and

$$\dot{\mathbf{U}} = \mathbf{Df} \cdot \mathbf{U} \tag{5.26}$$

as a coupled set of $2n$ differential equations. We recommend choosing a random initial ball for genericity.

2. At various times t_j , replace \mathbf{U} with the orthogonal axes of the ellipsoid defined by \mathbf{U} , using the Gram-Schmidt orthogonalization procedure. This can be done either every time T , for some suitable choice of T , or every time the integrator takes a step. We have found the latter prescription to be especially robust in practice.
3. If the length of any axis exceeds some *very* large value (say, near the maximum representable floating point value), normalize the ellipsoid and record the axis lengths

$$R_i^{(k)} \quad (\textit{ith axis at } k\textit{th rescaling}) \tag{5.27}$$

at the rescaling time. Do the same if any axis is smaller than some very small number.

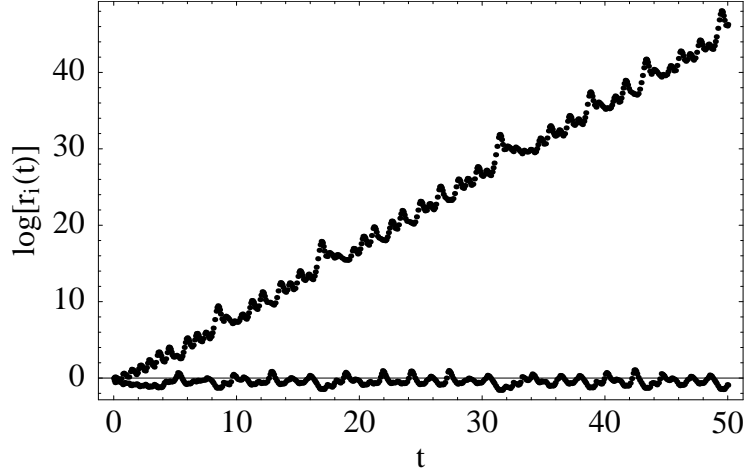


Figure 5.6: Closeup of Fig. 5.5, showing the natural logarithms of the two largest ellipsoid axes vs. t for the Lorenz system, calculated using the Jacobian method (Sec. 5.2.2). The slopes are the Lyapunov exponents. The plot for the larger axis closely matches the figures for the rescaled deviation vector method (Fig. 5.2) and the single tangent vector Jacobian method (Fig. 5.3).

4. Record the value of

$$\log r_i^{(j)} = \log [L_i(t_j)] + \sum_{k=1}^{k_{\max}} \log R_i^{(k)} \quad (5.28)$$

at each time t_j , where L_i is the i th principal axis length. The second term accounts for the axis lengths at the k_{\max} rescaling times. Note that if t_j is a rescaling time itself, then $\log [L_i(t_j)] = \log 1 = 0$, since by construction the ellipsoid has been normalized back to a unit ball.

5. After reaching the final number of time steps N , perform a least squares fit on the pairs $(t_j, \log r_i^{(j)})$ to find the slopes λ_i . Since

$$\log [r_i(t)] \approx \lambda_i t, \quad (5.29)$$

the slope λ_i is the Lyapunov exponent corresponding to the i th principal axis. Using the Gram-Schmidt procedure should result in the relationship $\lambda_1 > \dots > \lambda_n$.

Most of the value of calculating λ_i for $i > 1$ comes from having *all* n of the exponents (Sec. 5.2.3 below). Nevertheless, it is worth noting that the algorithm works for any value $0 < m \leq n$, so the method above can be used without alteration to find an arbitrary number of exponents. Fig. 5.6 shows the axis growth for $m = 2$ in the Lorenz system, while Fig. 5.5 shows the growth for $m = n = 3$.

5.2.3 The value of multiple exponents

Calculating all the exponents of a system of differential equations allows us to paint a more complete picture of the dynamics in several different ways. In particular, with all n exponents comes the ability to visualize the entire phase space ellipsoid (instead of just its principal axis), as in Fig. 5.4. Another important benefit of knowing all the exponents is a determination of dissipative or conservative behavior. Conservative flows preserve phase space volumes, while dissipative flows contract volumes. Geometrically, the volume V of an ellipsoid is proportional to the product of its principal axes $\{r_i\}$, so that the ratio of the final to the initial volume is

$$\frac{V_f}{V_0} = \prod_i r_i, \quad (5.30)$$

assuming that the initial volume is a unit ball. For dissipative systems, phase space volumes in general contract exponentially according to

$$\frac{V_f}{V_0} = e^{-\Lambda t}, \quad (5.31)$$

where Λ is a positive constant. Combining Eq. (5.30) and Eq. (5.31) yields

$$\Lambda = -\log\left(\frac{V_f}{V_0}\right) = -\log\left(\prod_i r_i\right) = -\sum_i \log r_i = -\sum_i \lambda_i, \quad (5.32)$$

where the λ_i are the Lyapunov exponents. In other words, *the phase space volume contraction constant Λ is equal to minus the sum of the Lyapunov exponents.*

If the Lyapunov exponents sum to zero, then the contraction factor vanishes, and volumes are conserved—i.e., the system is conservative. The special case of Hamiltonian systems is of particular interest, since the equations of motion for many mechanical systems can be derived from a Hamiltonian. The Hamiltonian property strongly constrains the Lyapunov exponents, which must *cancel pairwise*: to each exponent $+\lambda$ there corresponds a second exponent $-\lambda$ [11]. Several examples of this $\pm\lambda$ property of Hamiltonian systems appear below.

Having all the Lyapunov exponents also allows us to verify that there is at least one vanishing exponent, corresponding to motion tangent to the flow, which must be the case for any chaotic system. (See Ref. [9] for a proof.) Since we have finite numerical precision, we do not expect to find any exponent to be identically zero, but *some* exponent should always be close to zero. A practical criterion for “close to zero” is to compute error estimates for the least-squares fits advocated in Sec. 5.2.2.4; an exponent is “close to zero” if it is zero to within the standard error of the fit. Applications of this method appear in Sec. 5.2.4.1 and Sec. 5.2.4.2 below. It is worth noting that the fitting errors are not the dominant source of variance in calculating Lyapunov exponents; variations

in the initial conditions and initial deviation vectors contribute more to the uncertainty than errors in the fits. See Sec. 5.2.4.1 for further discussion.

One final note deserves mention: the statement that $\Lambda = -\sum_i \lambda_i$ is equivalent to a theorem due to Liouville [9], which relates the volume contraction to the trace of the Jacobian matrix:

$$\frac{V_f}{V_0} = \exp \left(\int_0^t \text{Tr} \mathbf{Df}(t) dt \right), \quad (5.33)$$

where again we assume that V_0 corresponds to a unit ball. If the trace of the Jacobian matrix happens to be time-independent, then this yields

$$\frac{V_f}{V_0} = \exp [(\text{Tr} \mathbf{Df}) t], \quad (\text{time-independent trace}) \quad (5.34)$$

so that Eq. (5.32) gives $\Lambda = -\text{Tr} \mathbf{Df}$. In this special case, we can perform a consistency check by verifying that

$$\sum_i \lambda_i = \text{Tr} \mathbf{Df}. \quad (\text{time-independent trace}) \quad (5.35)$$

5.2.4 Examples

5.2.4.1 The Lorenz system

Following the phase space ellipsoid allows us to visualize the dynamics of the Lorenz system in an unusual way. Fig. 5.4 shows the Lorenz attractor together with the phase space ellipsoid for a short amount of time ($t_f = 0.4$). The initial ball is evolved using Eq. 5.21, so it represents the true tangent space evolution, which is then superposed on the Lorenz phase space (x, y, z) . It is evident that the initial ball is stretched in one direction and flattened in another, as well as rotated. (As we shall see, the third direction is neither stretched nor squeezed, corresponding to the zero exponent discussed in Sec. 5.2.3.)

By recording natural logarithms of the ellipsoid axes as the system evolves, we can obtain numerical estimates for the Lyapunov exponents, as discussed in Sec. 5.2.2.4. A plot of $\log[r_i(t)]$ vs. t appears in Fig. 5.5 for a final time $t_f = 50$, with the slopes giving approximate values for the exponents. Using a $t_f = 5000$ integration for greater accuracy yields the estimates

$$\begin{aligned} \lambda_1 &= 0.905 \pm 9 \times 10^{-6} \\ \lambda_2 &= 1.5 \times 10^{-6} \pm 1.7 \times 10^{-6} \\ \lambda_3 &= -14.57 \pm 9 \times 10^{-6} \end{aligned} \quad (5.36)$$

for the parameter values $\sigma = 10$, $b = 8/3$, and $r = 28$. The \pm values are the standard errors on the least-squares fit of $\log[r_i(t)]$ vs. t . One of the exponents is close to zero (as required for a flow) in

the sense of Sec. 5.2.3: the error in the fit not small compared to the exponent. [In the case shown in Eq. (5.36), the “error” is actually *larger* than the exponent.] The other two exponents are clearly nonzero, with the positive exponent indicating chaos.

As mentioned briefly in Sec. 5.2.3, the largest source of variance in calculating Lyapunov exponents is variations in the initial conditions, not errors in the least-squares fits used to determine the exponents. We express the exponents in the form

$$\bar{\lambda} \pm \frac{\sigma}{\sqrt{N}}, \quad (5.37)$$

where

$$\bar{\lambda} = \frac{1}{N} \sum_{j=1}^N \lambda^{(j)} \quad (5.38)$$

is the sample mean and

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (\lambda^{(j)} - \bar{\lambda})^2} \quad (5.39)$$

is the standard deviation. For the Lorenz system, using a final integration time of $t_f = 5000$ for $N = 50$ random initial balls [all centered on the same initial value of (x_0, y_0, z_0)] gives

$$\begin{aligned} \lambda_1 &= 0.9053 \pm 4.1 \times 10^{-4} \\ \lambda_2 &= -4.5 \times 10^{-6} \pm 7.6 \times 10^{-7} \\ \lambda_3 &= -14.5720 \pm 4.1 \times 10^{-4} \end{aligned} \quad (5.40)$$

The values of the error are much greater than the standard errors associated with the least-squares fit for the slope for any one trial. As expected, it is evident that λ_2 is consistent with zero.

There is a strongly expanding direction and a very strongly contracting direction in the Lorenz system, and the volume contraction constant Λ is large: $\Lambda = -\sum_i \lambda_i = 13.67$, so that after a time $t = 5000$ the volume is an astonishingly small 6.75×10^{-29674} . This is despite the exponential growth of the largest principal axis, which grows in this same time to a length 1.52×10^{1965} ; the volume nevertheless contracts, since the smallest axis shrinks to 4.44×10^{-31639} in the same time. We note that the periodic renormalization and reorthogonalization of the ellipsoid axes is absolutely essential from a numerical perspective, since these axis lengths are far above and below the floating point (double precision) limits of $\mathbf{xmax} \approx \mathbf{xmin}^{-1} \approx 10^{308}$ on a typical IEEE-compliant machine [12].

The Lorenz system affords an additional check on the numerically determined exponents: the trace of the Jacobian matrix [Eq. (5.13)] is time-independent, so the exponents should satisfy Eq. (5.35):

$$\sum_i \lambda_i = -13.67 \stackrel{?}{=} \text{Tr } \mathbf{Df} = -(\sigma + 1 + b) = -\frac{41}{3} \approx -13.67. \quad (5.41)$$

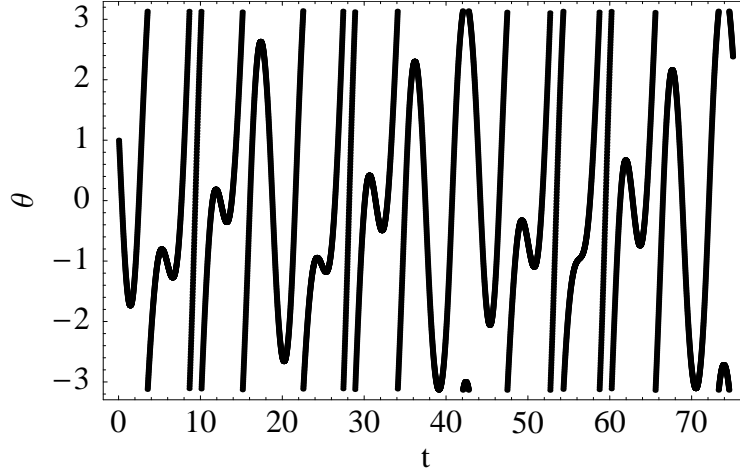


Figure 5.7: θ vs. t for the forced damped pendulum [Eq. (5.42)].

Eq. (5.35) is thus well satisfied.

5.2.4.2 The forced damped pendulum

We turn now to our second principal example of a chaotic dynamical system, the forced damped pendulum (FDP). This is a standard pendulum with damping and periodic forcing; written as a first-order ODE, our equations are as follows:

$$\begin{aligned}\dot{\theta} &= \omega \\ \dot{\omega} &= -c\omega - \sin\theta + \rho \sin t \\ \dot{t} &= 1\end{aligned}\tag{5.42}$$

Here c is the damping coefficient and ρ is the forcing amplitude, and the gravitational acceleration g and pendulum length ℓ are set to one for simplicity. We include the equation $\dot{t} = 1$ so that the system is autonomous (i.e., we remove the explicit time-dependence by treating time as a dynamical variable with unit time derivative). In addition to being an example with transparent physical relevance (in contrast to the Lorenz system), the forced damped pendulum, in slightly altered form, serves as a model constrained system in Sec. 5.3 below.

The forced damped pendulum is chaotic for many values of c and ρ . For simplicity, in the present case we fix $c = 0.1$ and $\rho = 2.5$. A plot of θ vs. t shows the system's erratic behavior (Fig. 5.7), but a more compelling picture of the dynamics comes from a time- 2π stroboscopic map. A time- T map involves taking a snapshot of the system every time T and then plotting ω vs. θ . Since the forcing term in Eq. (5.42) is 2π -periodic, this provides a natural value for T in the present case. The resulting plot shows the characteristic folding and stretching of a fractal attractor (Fig. 5.8), which

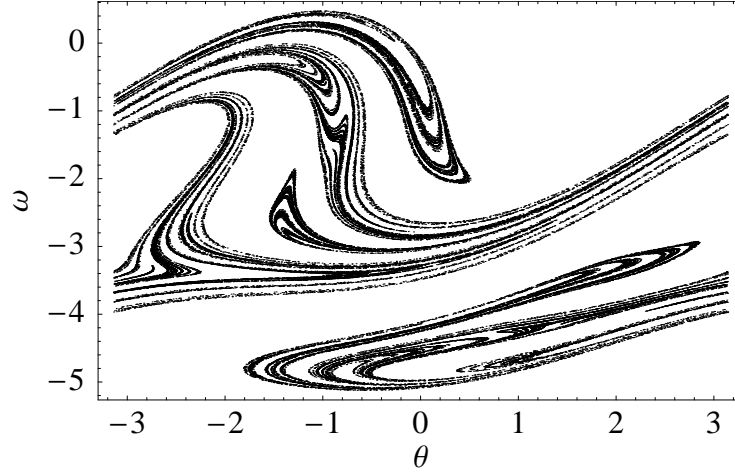


Figure 5.8: ω vs. θ : the time- 2π stroboscopic map for the forced damped pendulum. A point $(\omega = \dot{\theta}, \theta)$ is plotted every time 2π , resulting in a fractal attractor characteristic of dissipative chaos.

for the FDP attracts almost all initial conditions [9].

The forced damped pendulum is dissipative and strongly chaotic. We calculate the Lyapunov exponents (Fig. 5.9) using the Jacobian matrix:

$$\mathbf{Df} = \begin{pmatrix} 0 & 1 & 0 \\ -\cos \theta & -c & \rho \cos t \\ 0 & 0 & 0 \end{pmatrix}, \quad (5.43)$$

The Lyapunov exponents are (for a $t_f = 5 \times 10^4$ integration)

$$\begin{aligned} \lambda_1 &= 0.160 \pm 7 \times 10^{-6} \\ \lambda_2 &= 8 \times 10^{-8} \pm 1 \times 10^{-7} \\ \lambda_3 &= -0.262 \pm 7 \times 10^{-6} \end{aligned} \quad (5.44)$$

where the error terms are the standard errors in the least-squares fit for the slope. (See Sec. 5.4 and especially Table 5.1 for the true errors due to varying initial deviations.) One exponent is consistent with zero (as required for a flow) to within the error of the fit. The dissipation constant is $\Lambda = -\sum_i \lambda_i = 0.1$. The trace of the Jacobian matrix is time-independent, so that $\text{Tr } \mathbf{Df} = -c$, and indeed $\sum_i \lambda_i = -0.1 = -c = \text{Tr } \mathbf{Df}$ as predicted by Eq. 5.35.

The zero exponent in the FDP is associated with the time “degree of freedom” in the Jacobian: if we delete the final row and column of the Jacobian matrix, only the positive and negative exponents remain (see, e.g., Fig. 5.13 below). Since the time is not an actual dynamical variable, for the remainder of this chapter we will suppress this “time piece,” but it is important to note that the

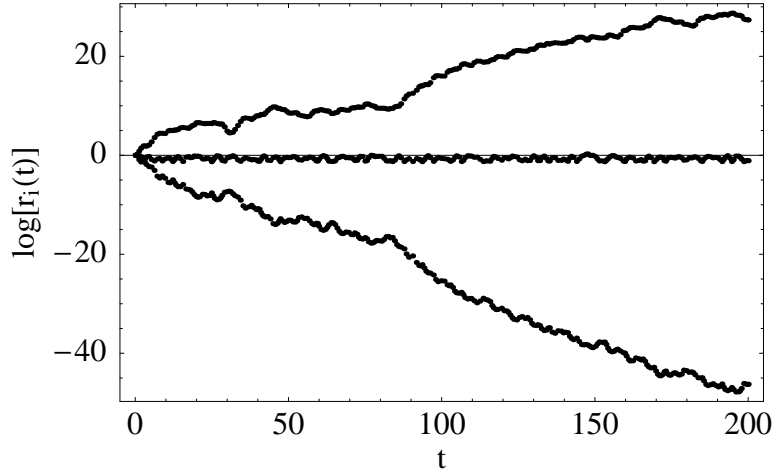


Figure 5.9: The natural logarithms of all three of the ellipsoid axes r_i vs. t for the forced damped pendulum, calculated using the Jacobian method (Sec. 5.2.2). The slopes are the Lyapunov exponents. The three lines correspond to the exponents $\lambda_1 = 0.160 \pm 0.0049$, $\lambda_2 = 0.0$, and $\lambda_3 = -0.262 \pm 0.0053$ (Sec. 5.4 and Table 5.1).

time dependence is absolutely crucial to the presence of chaos. According to the *Poincaré-Bendixon theorem* [9], an autonomous system of differential equations with fewer than three degrees of freedom *cannot* be chaotic. We will treat the FDP system as a time-dependent system with two degrees of freedom, but the extra equation $\dot{t} = 1$ in the autonomous formulation is what creates the potential for chaos.

An instructive case to consider is the limit $c = \rho = 0$. In this limit, the system is a simple pendulum, which is a Hamiltonian system. A simple pendulum is not chaotic, of course, and both its Lyapunov exponents are zero, but the Hamiltonian character of the system nevertheless shows up in the $\pm\lambda$ property discussed above (Sec. 5.2.3): numerically, the exponents approach zero in a symmetric fashion, as shown in Fig. 5.10.

5.3 Lyapunov exponents in constrained flows

We come now to the *raison d'être* of this chapter, namely, the calculation of Lyapunov exponents for constrained systems. For pedagogical purposes, our primary example is the forced damped pendulum with the position written in Cartesian coordinates. In addition to this instructive example, we also discuss two constrained systems of astrophysical interest, involving the orbits of spinning compact objects such as neutron stars or black holes (see, e.g., Chapters 2 and 3 and references therein).

Written in terms of the Cartesian coordinates $(x, y) = (\cos \theta, \sin \theta)$, the equations of motion for

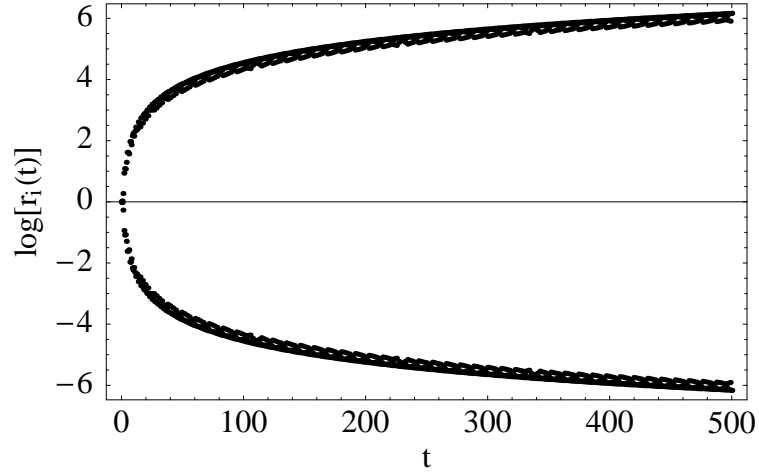


Figure 5.10: The natural logarithms of the ellipsoid axes r_i vs. t for the forced damped pendulum in the limit of zero dissipation and zero forcing (i.e., a simple pendulum). The Lyapunov exponents are zero, and the distance between nearby trajectories grows linearly (leading to logarithmic growth in this log plot). Nevertheless, the Hamiltonian character of the system is manifest in the $\pm\lambda$ symmetry: for each exponent $+\lambda$, there is a corresponding exponent $-\lambda$. In the nonchaotic limiting case shown here, the Lyapunov exponents approach zero symmetrically.

the FDP [Eq. (5.42)] become (upon suppressing the time piece)

$$\begin{aligned}\dot{x} &= -\omega y \\ \dot{y} &= \omega x \\ \dot{\omega} &= -c\omega - y + \rho \sin t\end{aligned}\tag{5.45}$$

For a pendulum with unit radius, the Cartesian coordinates of the pendulum satisfy the constraint

$$x^2 + y^2 = 1.\tag{5.46}$$

Although it is certainly possible to use (\dot{x}, \dot{y}) in the equations of motion, along with (x, y) , this is an unnecessary complication; in order to keep the equations as simple as possible, we retain the variable ω in the equations of motion.

Developing the techniques for solving constrained systems using this toy example has several advantages. The equations of motion and the constraint are extremely simple, which makes it easy to see the differences between the constrained and unconstrained cases. In addition, the constraint is easy to visualize, and yet it captures the key properties of much more complicated constraints. Finally, since we have already solved the same problem in unconstrained form, it is easy to verify that the techniques of this section reproduce the results from Sec. 5.2.4.2.

5.3.1 Constraint complications

To see how constraints complicate the calculation of Lyapunov exponents, consider an implementation of the deviation vector approach (Sec. 5.2.1). In the unconstrained forced damped pendulum, given an initial condition, we would construct a deviated trajectory separated by a small angle $\delta\theta$ (and a small velocity $\delta\omega$). In the constrained version, a naïve implementation would use a deviated trajectory with spatial coordinates $x + \delta x$ and $y + \delta y$, where $\delta\mathbf{y} = (\delta x, \delta y)$ is a small but otherwise arbitrary deviation vector. But the deviations are not independent; the deviated initial condition must satisfy the constraint:

$$(x + \delta x)^2 + (y + \delta y)^2 = 1. \quad (5.47)$$

To lowest order in δx , we must have $\delta y = -(x/y)\delta x$.

We can now consider a more general case. Suppose there are k constraints, which we write as a k -dimensional vector equation $\mathbf{C}(\mathbf{y}) = \mathbf{0}$. (In our example, \mathbf{C} has only one component: with $\mathbf{y} = (x, y, \omega)$, we have $C_1(\mathbf{y}) = x^2 + y^2 - 1 = 0$.) Then if a point \mathbf{y} satisfies the constraints, the deviated trajectory must satisfy them as well:

$$\mathbf{C}(\mathbf{y} + \delta\mathbf{y}) = \mathbf{0}. \quad (5.48)$$

We will refer such a $\delta\mathbf{y}$ as a *constraint-satisfying deviation*.

Let us outline one possible method for constructing such a constraint-satisfying deviation. Let n be the number of phase space coordinates ($n = 3$ for the constrained forced damped pendulum model). Consider an n -dimensional vector $\tilde{\mathbf{y}}_0$ that has d nonzero entries, where d represents the true number of degrees of freedom ($d = 2$ for the constrained FDP). Assume that we have some method for constructing from $\tilde{\mathbf{y}}_0$ an n -dimensional initial condition \mathbf{y}_0 that satisfies the constraints. For example, we could specify the initial values of x and ω , and then derive an initial value of y using $y = \sqrt{1 - x^2}$ (or $y = -\sqrt{1 - x^2}$; more on this later). Now consider an n -dimensional vector $\tilde{\mathbf{y}}'_0 = \tilde{\mathbf{y}}_0 + \delta\tilde{\mathbf{y}}_0$, which adds *arbitrary* deviations to d degrees of freedom. We can then use the same method as above to find \mathbf{y}'_0 from $\tilde{\mathbf{y}}'_0$, and then set

$$\delta\mathbf{y}_0 = \mathbf{y}'_0 - \mathbf{y}_0 \quad (5.49)$$

to arrive at a constraint-satisfying deviation.

5.3.2 Constrained deviation vectors

Having determined $\delta\mathbf{y}_0$ by Eq. (5.49) (or by some other method), we can immediately apply the unrescaled deviation vector approach: simply track \mathbf{y}' and \mathbf{y} as the two trajectories evolve, and

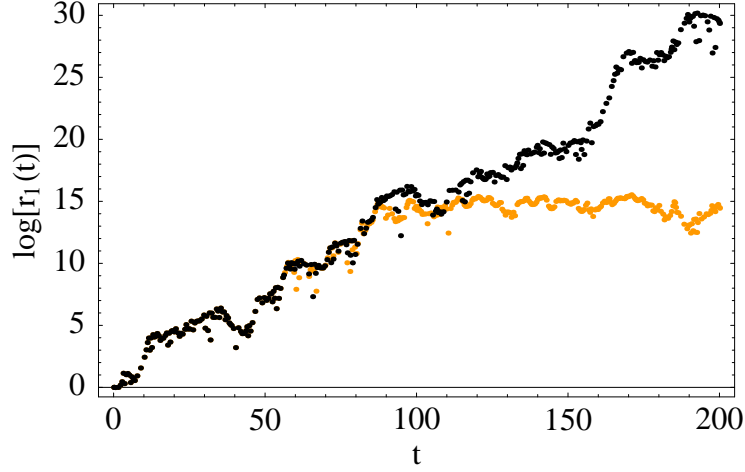


Figure 5.11: Comparison of the unrescaled (light) and rescaled (dark) constrained deviation vector methods for calculating the principal Lyapunov exponent of the constrained forced damped pendulum (Sec. 5.3.2.1). The slope of the rescaled line is the Lyapunov exponent, $\lambda_1 = 0.161 \pm 0.0046$ (Sec. 5.4). The initial deviation is $\|\delta\mathbf{y}_0\| = 10^{-6}$, and rescaling occurs (for the rescaled method) if $\|\delta\mathbf{y}\| \geq 10^{-2}$, which happens 4 times in this figure. As in Fig. 5.2, the unrescaled approach saturates once the deviation has grown too large.

monitor the length of $\delta\mathbf{y} = \mathbf{y}' - \mathbf{y}$. Since the equations of motion preserve the constraint, the resulting $\delta\mathbf{y}$ is always constraint-satisfying. The only subtlety is using a restricted norm to eliminate the extra degrees of freedom; for example, the restricted FDP norm is

$$\|\delta\mathbf{y}\|_r = \sqrt{\delta x^2 + \delta\omega^2} \quad (5.50)$$

if we choose to eliminate the y degree of freedom. Since $\delta y \approx -(x/y)\delta x$, using the full Euclidean distance would add the term $\delta y^2 = (x^2/y^2)\delta x^2$ to the expression under the square root, leading to an overestimate for the principal exponent. The restricted norm avoids this problem by considering only true degrees of freedom.

5.3.2.1 Rescaling for constrained systems

In contrast to the simplicity of the unrescaled method, the *rescaled* deviation vector method requires great care, since a carelessly rescaled deviation is not constraint-satisfying: $\mathbf{C}(\mathbf{y} + \delta\mathbf{y}/r) \neq \mathbf{0}$ for a rescaling factor $r \neq 1$. In this case, it is necessary to extract $\delta\tilde{\mathbf{y}}$ from $\delta\mathbf{y}$ and then rescale it back to its initial size $\|\delta\tilde{\mathbf{y}}_0\|$ using the restricted norm. By reapplying the procedure leading to Eq. (5.49), we then find a new (rescaled) constraint-satisfying $\delta\mathbf{y}$ that satisfies $\|\delta\mathbf{y}\|_r = \|\delta\tilde{\mathbf{y}}_0\|$. In this case, *it is essential that the new deviation vector have the same constraint branches as the old one*. For example, suppose that in the FDP case the value of y is negative before the rescaling. When calculating a new \mathbf{y}' to arrive at the rescaled deviation $\delta\mathbf{y}$, it is then essential to choose

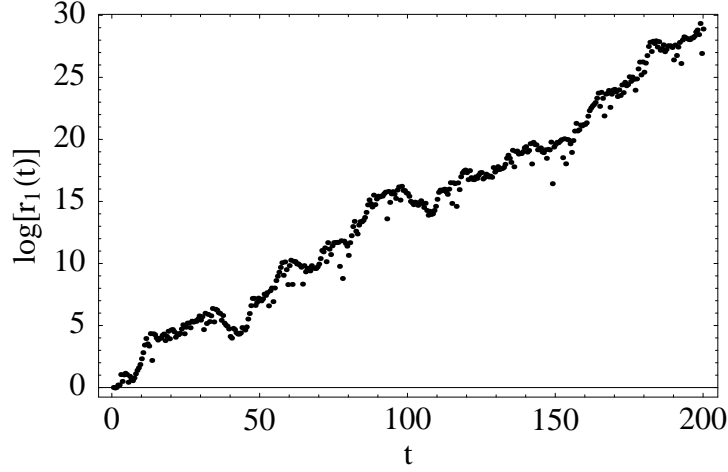


Figure 5.12: The natural logarithm of the tangent vector length $r_1 \equiv \|\xi(t)\|_r$ vs. t for the constrained forced damped pendulum, using a constraint-satisfying tangent vector (Sec. 5.3.2.2). We use the restricted norm $\|\cdot\|_r$ to calculate phase space distances (see text). Compare to Fig. 5.9 (unconstrained Jacobian method) and Fig. 5.11 (constrained deviation vector method).

the negative branch in the equation $y' = \pm\sqrt{1-x'^2}$. The result of implementing this constrained deviation vector method to the forced damped pendulum appears in Fig. 5.11.

5.3.2.2 A Jacobian method for the largest exponent

The method outlined above for unrescaled deviation vectors leads to a remarkably simple implementation of the single tangent vector Jacobian method. Given a constraint-satisfying deviation $\delta\mathbf{y}_0$, set

$$\xi_0 = \delta\mathbf{y}_0 / \|\delta\mathbf{y}_0\|_r, \quad (5.51)$$

where $\|\cdot\|_r$ is a restricted norm on the d true degrees of freedom. We refer to such a ξ as a *constraint-satisfying tangent vector*. Since the equations of motion preserve the constraints, we can evolve this tangent vector using Eq. (5.17). The Jacobian method does not saturate, so we need only rescale if $\|\xi\|_r$ approaches the floating point limit of the computer. We can then use a procedure based on the rescaled deviation method to find a new (rescaled) constraint-satisfying tangent vector, but this is typically unnecessary since by the time the floating point limit has been reached we already have a good estimate of the principal Lyapunov exponent. The resulting Lyapunov plot for the constrained FDP appears in Fig. 5.12.

5.3.2.3 Ellipsoid constraint complications

We now have three methods at our disposal for calculating the *largest* Lyapunov exponent, but for d degrees of freedom there are d exponents. What of these other exponents? Here we find an essential

difficulty in implementing the ellipsoid method described in Sec. 5.2.2.3. The core problem is this: the tangent vectors must be orthogonalized in order to extract all d principal ellipsoid axes, but at the same time each tangent vector must be constraint-satisfying. Simply put, it is impossible in general to satisfy the requirements of orthogonality and constraint satisfaction simultaneously.

We present here two different solutions to this problem, which we will refer to as the restricted Jacobian method and the constrained ellipsoid method.

5.3.3 Restricted Jacobian method

The most natural response to a system with more coordinates n than degrees of freedom d is to eliminate the spurious degrees of freedom using the constraints. Unfortunately, this procedure is often difficult in practice: solving the constraint equations may involve polynomial or transcendental equations that have no simple closed form. Even for the simple case of the FDP, the sign ambiguity in $y = \pm\sqrt{1-x^2}$ makes a simple variable substitution impractical. Fortunately, such substitutions are unnecessary: since the equations of motion preserve the constraints, there is no need in general to eliminate $n - d$ coordinates. In fact, constraints can be a virtue, since they can be used to check the accuracy of the integration.

The same cannot be said of the Jacobian matrix. As argued above, the extra degrees of freedom lead to fundamental difficulties in applying the Jacobian method for finding Lyapunov exponents; constraints, far from being a virtue, are a considerable complication. In contrast to the equations of motion, though, it is relatively straightforward to eliminate the spurious degrees of freedom. The trick is to write a *restricted* $d \times d$ Jacobian matrix, with entries only for d coordinates.

An example should make this clear. For the FDP system in constrained form, we wish to eliminate one degree of freedom in the Jacobian matrix, and we can choose to eliminate either x or y . Choosing the latter, the Jacobian becomes

$$\mathbf{Df} = \begin{pmatrix} \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial \omega} \\ \frac{\partial \dot{\omega}}{\partial x} & \frac{\partial \dot{\omega}}{\partial \omega} \end{pmatrix}, \quad (5.52)$$

where we have suppressed the derivatives with respect to the “time degree of freedom” (as discussed in Sec. 5.2.4.2). The term to focus on here is $\partial \dot{x} / \partial x$, which seems to be zero *a priori* since $\dot{x} = -\omega y$, but this is only true if we treat x and y as independent. Since we are eliminating the y degree of freedom, we *cannot* treat them as independent; y has a nonzero derivative with respect to x , so that

$$\frac{\partial \dot{x}}{\partial x} = -\omega \frac{\partial y}{\partial x}. \quad (5.53)$$

If we find $\partial y / \partial x$ using $y = \pm\sqrt{1-x^2}$, we have exactly the same sign ambiguity problem that we had in trying to eliminate the y degree of freedom in the equations of motion. The difference here

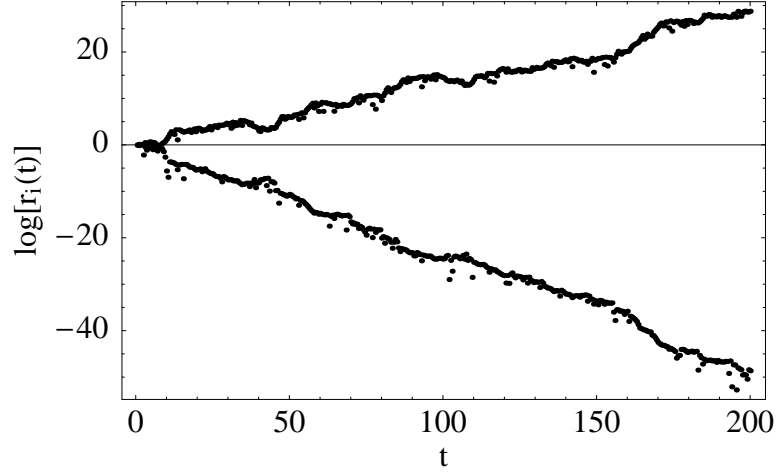


Figure 5.13: The natural logarithms of both ellipsoid axes for the constrained forced damped pendulum, calculated using the restricted Jacobian method (Sec. 5.3.3). The slopes are the Lyapunov exponents. The results agree well with the unconstrained case (Fig. 5.9 and Table 5.1).

is that we need only the *derivative* of y , not an explicit solution for y in terms of x , and this we can achieve by differentiating the constraint:

$$0 = \frac{\partial}{\partial x}(x^2 + y^2) = 2x + 2y \frac{\partial y}{\partial x} \Rightarrow \frac{\partial y}{\partial x} = -\frac{x}{y}. \quad (5.54)$$

If we integrate the equations of motion using the variables (x, y, ω) , then we have the value of y at any particular time, and we never need deal with the sign ambiguity. Using the same trick to calculate $\partial \dot{\omega} / \partial x$, we can write the *restricted Jacobian* as

$$\mathbf{Df} = \begin{pmatrix} \omega \frac{x}{y} & -y \\ \frac{x}{y} & -c \end{pmatrix} \quad (5.55)$$

We now proceed exactly as in the unconstrained Jacobian method, using the restricted Jacobian to calculate the evolution of the initial tangent space ball. Since we deal only with a number of coordinates equal to the true number of degrees of freedom, the constraints are not a consideration, and we can reorthogonalize exactly as before.

The general case is virtually the same. For n coordinates and d degrees of freedom, there must be $m = n - d$ constraint equations of the form

$$C_k(\mathbf{y}) = 0 \quad (5.56)$$

for $k = 1 \dots m$. We must choose which d coordinates to keep in the Jacobian matrix, eliminating m

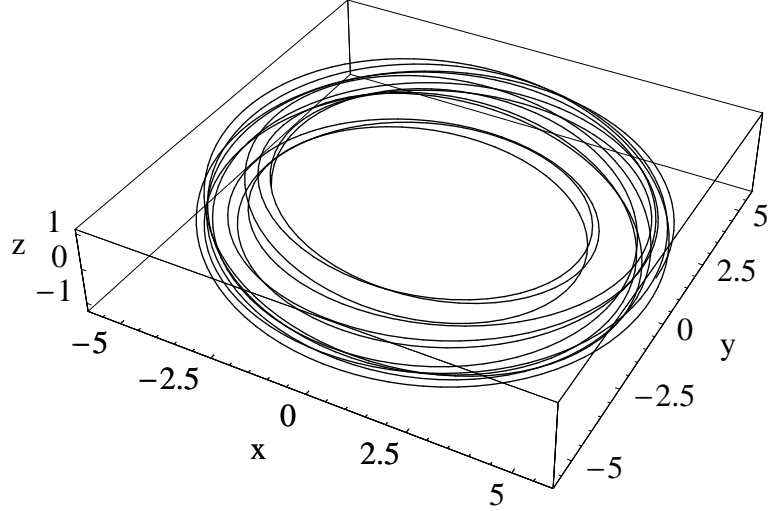


Figure 5.14: The orbit of a spinning relativistic binary, calculated using the post-Newtonian equations of motion. The equations model *two* spinning bodies, but we use an effective one-body approach to reduce the dynamics to the motion of one body. Distances are measured in terms of GM/c^2 , where $M = m_1 + m_2$ is the total mass of the system. For a pair of black holes, each with 10 times the mass of the Sun, the length unit is $GM/c^2 = 20 GM_\odot/c^2 = 30$ km.

coordinates in the process. By differentiating the constraints, we arrive at m *linear* equations for the derivatives of the m eliminated coordinates in terms of the n variables:

$$\frac{\partial C_k}{\partial y_j} = 0, \quad (5.57)$$

where j ranges over the indices of the eliminated coordinates ($j = 2$, corresponding to y , for the FDP). Since these are linear equations, they are both easy to solve and do not suffer from any sign or branch ambiguities. The $d \times d$ Jacobian matrices that result allow the calculation of Lyapunov exponents with all the robustness of the Jacobian method for unconstrained systems.

We considered the constrained forced damped pendulum for purposes of illustration, but it is admittedly artificial. A more realistic example is shown in Fig. 5.14, which illustrates the dynamics of two spinning black holes with comparable masses. (Such systems are of considerable interest for ground-based gravitational wave detectors such as the LIGO project.) The equations of motion come from the Post-Newtonian (PN) expansion of full general relativity—essentially, a series expansion in the dimensionless velocity v/c , where the first term is ordinary Newtonian gravity and the higher-order terms are post-Newtonian corrections (see, e.g., [14–16]). The constraint comes from the spins of the black holes: it is most natural to think of the spin as having *two* degrees of freedom (a fixed magnitude with two variable angles specifying the location on a sphere), but the equations of motion use all *three* components of each hole’s spin. We apply the methods described above to eliminate

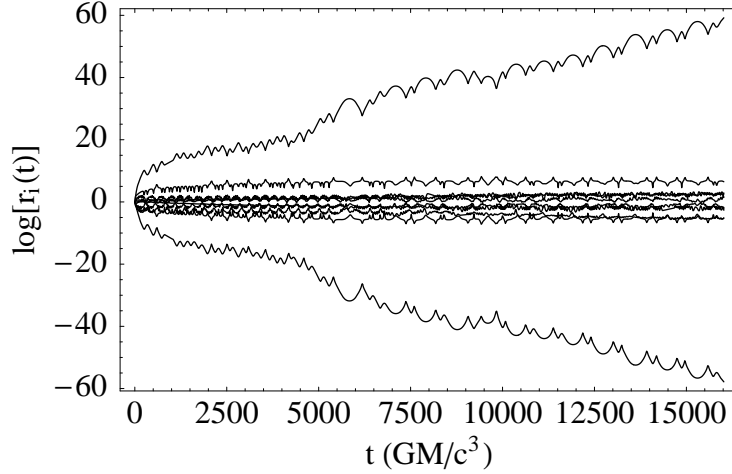


Figure 5.15: The natural logarithms of the ellipsoid axes r_i vs. t for the system shown in Fig. 5.14. Time is measured in units of GM/c^3 , where $M = m_1 + m_2$ is the total mass of the system. For two 10 solar-mass black holes, the time unit is $GM/c^3 = 20 GM_\odot/c^3 = 10^{-4}$ s. The spin magnitudes are fixed, so that each spin vector represents only two true degrees of freedom. We deal with this constraint by using the restricted Jacobian method (Sec. 5.3.3). Two nonzero exponents are clearly visible, but all the others are consistent with zero. Note the $\pm\lambda$ symmetry characteristic of Hamiltonian systems.

one of the spin degrees of freedom for each black hole, using the constraints

$$S_{x,i}^2 + S_{y,i}^2 + S_{z,i}^2 = S_i^2 = \text{const.}, \quad i \in \{1, 2\}. \quad (5.58)$$

Using the *effective one-body* approach [14], *a priori* the system has 12 degrees of freedom: three each for relative position \mathbf{x} , momentum \mathbf{p} , and the spins \mathbf{S}_1 and \mathbf{S}_2 . Eliminating two spin components leaves 10 true degrees of freedom. As a result, the system has 10 Lyapunov exponents, as shown in Fig. 5.15; note in particular the $\pm\lambda$ symmetry characteristic of Hamiltonian systems.

5.3.4 Constrained ellipsoid method

The restricted Jacobian method relies on eliminating spurious degrees of freedom from the Jacobian matrix, but such a prescription relies on making a choice—namely, which coordinates to eliminate. Each choice results in a different Jacobian matrix. Since calculating the Jacobian matrix even once can be a formidable task for sufficiently complicated systems, it is valuable to have a method that uses the *full* Jacobian—treating all coordinates as independent—which can be calculated once and then never touched again. This requirement leads to the *constrained ellipsoid method*, which uses the full Jacobian matrix to evolve *constraint-satisfying* tangent vectors, collectively referred to as a “constrained ellipsoid.” When recording ellipsoid axis growth, we extract from each vector a number of components equal to the true number of degrees of freedom, resulting in vectors that can

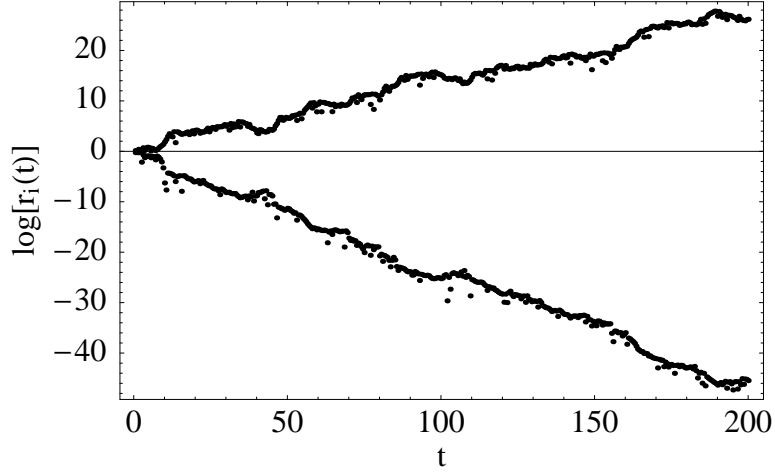


Figure 5.16: The natural logarithms of both ellipsoid axes r_i vs. t for the constrained forced damped pendulum, calculated using the constrained ellipsoid method (Sec. 5.3.4). The slopes are the Lyapunov exponents. The results agree well with the unconstrained case (Fig. 5.9 and Table 5.1).

be orthogonalized and (if necessary) normalized just as in the unconstrained case.

A detailed description of the constrained ellipsoid algorithm appears below, but we first present an important prerequisite: calculating constraint-satisfying tangent vectors. Let a tilde denote a vector with dimension d equal to the true number of degrees of freedom (as in Sec. 5.3.1). We construct a full tangent vector ξ (with n components) from a d -dimensional vector $\tilde{\xi}$ at a point \mathbf{y} on the flow as follows:

1. Let $\tilde{\mathbf{y}}' = \tilde{\mathbf{y}} + \epsilon \tilde{\xi}$ for a suitable choice of ϵ .
2. Fill in the missing components of $\tilde{\mathbf{y}}'$ using the constraints to form \mathbf{y}' as in Sec. 5.3.1.
3. Infer the full tangent vector ξ using

$$\xi = \frac{\mathbf{y}' - \mathbf{y}}{\epsilon}. \quad (5.59)$$

Setting the initial conditions is now simple: form a random $d \times d$ matrix, orthonormalize it, and then infer the full $d \times n$ matrix using the method above on each column. The construction of constraint-satisfying tangent vectors described above is also necessary in the reorthogonalization steps of the constrained ellipsoid method.

The full method is an adaptation of the Jacobian method from Sec. 5.2.2.4:

1. Construct a random $d \times d$ matrix and orthonormalize it to form a unit ball. Use the constraints to infer the full $d \times n$ matrix \mathbf{U} .
2. Evolve the system forward using the equations of motion and the evolution equation for \mathbf{U} ,

$$\dot{\mathbf{U}} = \mathbf{Df} \cdot \mathbf{U}. \quad (5.60)$$

3. At each time T , extract the relevant eight components from each tangent vector to form a $d \times d$ ellipsoid, orthonormalize it, and then fill in the missing components using the constraints, yielding again a $d \times n$ matrix. The restricted norms of the d tangent vectors contribute to the running sum for the logs of the ellipsoid axes [Eq. (5.28)].

It is important to note that, unlike the other Jacobian methods, rescaling every time time T (or some similar method) is required for the inference equation [Eq. (5.59)], since the product of ϵ and the components of ξ must be small for the inference to work correctly. The method only works if the system is renormalized regularly, so the value of T should be chosen to be small enough that no principal ellipsoid axis grows too large.

As before, we use the constrained FDP model for purposes of illustration. Treating each coordinate as independent yields [upon differentiation of Eq. (5.45)]:

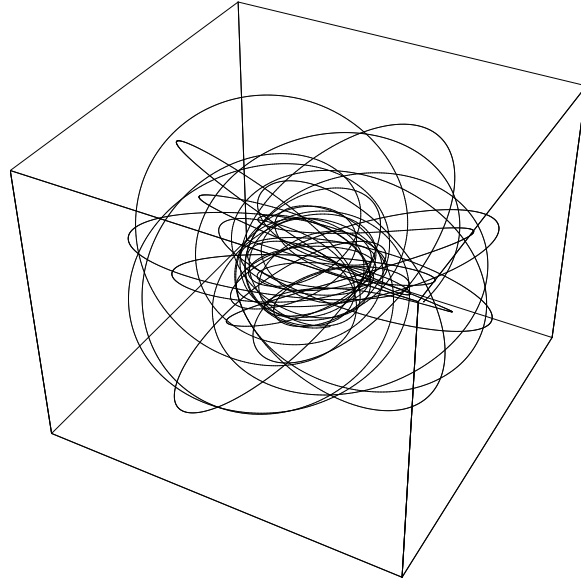
$$\mathbf{Df} = \begin{pmatrix} \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial y} & \frac{\partial \dot{x}}{\partial \omega} \\ \frac{\partial \dot{y}}{\partial x} & \frac{\partial \dot{y}}{\partial y} & \frac{\partial \dot{y}}{\partial \omega} \\ \frac{\partial \dot{\omega}}{\partial x} & \frac{\partial \dot{\omega}}{\partial y} & \frac{\partial \dot{\omega}}{\partial \omega} \end{pmatrix} = \begin{pmatrix} 0 & -\omega & -y \\ \omega & 0 & x \\ 0 & -1 & -c \end{pmatrix} \quad (5.61)$$

The coordinates are not independent, of course, but this Jacobian matrix satisfies Eq. (5.11) as long as the deviation is constraint-satisfying. For example, using the full deviation vector $\delta \mathbf{y} = (\delta x, \delta y, \delta \omega)$ with Eq. (5.61) gives the same result as the restricted deviation vector $\delta \tilde{\mathbf{y}} = (\delta x, \delta \omega)$ with Eq. (5.55), as long as $\delta y = -(x/y) \delta x$. As a result, the Lyapunov exponents calculated with the constrained ellipsoid method (Fig. 5.16) agree closely with the restricted Jacobian method (and with the original unconstrained results [Fig. (5.9)]).

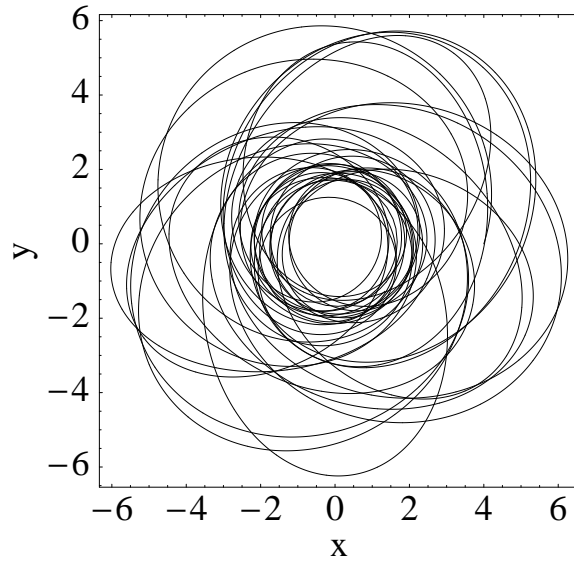
As a final example of the constrained ellipsoid method, consider Fig. 5.17, which shows a solution to equations that model a relativistic spinning test particle (e.g., a black hole or neutron star) orbiting a supermassive rotating black hole. (The case illustrated is a limiting case of the equations, which is mathematically valid but not physically realizable; see Chapter 2.) These equations (usually called the Papapetrou equations) are highly constrained, so a naïve calculation of the Lyapunov exponents is not correct. It was the complicated nature of the Jacobian matrix for this system that originally motivated the development of the methods in this section (Chapter 2). A Lyapunov plot corresponding to the orbit in Fig. 5.17 is shown in Fig. 5.18. Note especially the $\pm \lambda$ symmetry, a result of the Hamiltonian nature of the equations of motion.

5.4 Comparing the methods

A summary plot of all the methods discussed in this chapter, applied to the forced damped pendulum, appears in Fig. 5.19. It is evident that all the methods agree closely. A more quantitative comparison



(a)



(b)

Figure 5.17: The orbit of a small spinning compact object (such as a solar-mass black hole) in the spacetime of a rotating supermassive black hole. (a) The orbit embedded in spherical polar coordinates; (b) the orbit's projection onto the x - y plane. The lengths are expressed in terms of GM/c^2 , where M is the mass of the central black hole. For a maximally spinning black hole, the horizon radius is $r_H = GM/c^2$. For the supermassive black hole at the center of the Milky Way, $M = 3 \times 10^6 M_\odot$ [17], which corresponds to a length unit of $GM/c^2 = 4.4 \times 10^9$ m. The system shown here is chaotic (Fig. 5.18), although this orbit represents a limiting case of the equations that is not physically realizable, as shown in Chapter 2.

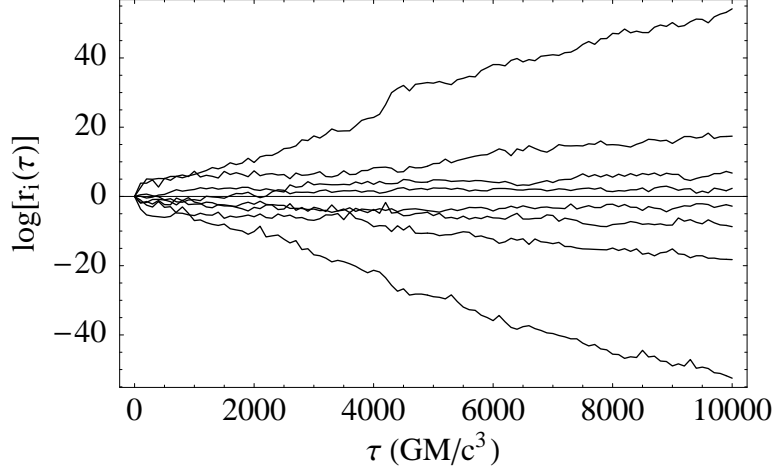


Figure 5.18: The natural logarithms of the ellipsoid axes for the system shown in Fig. 5.17 vs. relativistic proper time τ , in units of GM/c^3 , where M is the black hole's mass. For the supermassive black hole at the center of the Milky Way, $M = 3 \times 10^6 M_\odot$ [17], which corresponds to a time unit of $GM/c^3 = 15$ s. The largest Lyapunov exponent is $\lambda_{\max} \approx 5 \times 10^{-3} (GM/c^3)^{-1}$, which corresponds to an e -folding timescale of $\tau_\lambda = 1/\lambda = 2 \times 10^2 GM/c^3$. For $M = 3 \times 10^6 M_\odot$, this means that nearby trajectories diverge by a factor of e in the local (Lorentz) frame of an observer on this orbit in a time $\tau = 3000$ s = 50 min. We find nonzero exponents in this system only for physically unrealistic values of the small body's spin (Chapter 3).

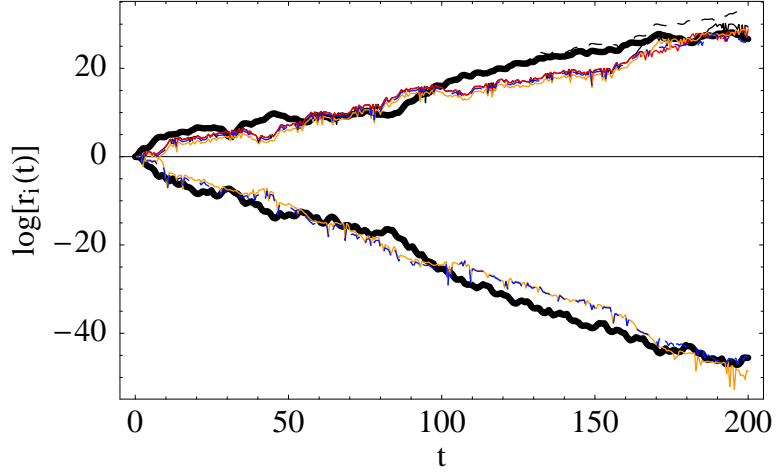


Figure 5.19: Natural logarithms of the ellipsoid axes vs. t for the unconstrained deviation vector method (dashed), the unconstrained Jacobian method from Fig. 5.9 (thick) and *all* the constrained methods. The constrained methods include the following: rescaled deviation vector (black), Jacobian with single constraint-satisfying tangent vector (red), restricted Jacobian (orange), and constrained ellipsoid (dashed blue). (The colors appear as shades of gray in print versions of this chapter.) All the constrained methods start with exactly the same initial conditions.

Table 5.1: Comparison of different Lyapunov exponent methods applied to the forced damped pendulum. We consider both the unconstrained [Eq. (5.42)] and constrained [Eq. (5.45)] formulations. The integrations have a final time $t_f = 10^4$, and for each method we consider 100 random initial deviations. We calculate the positive exponent (λ_1) and, if possible, the negative exponent (λ_3) as well. (We omit the zero exponent (λ_2) for brevity.) The error estimates are the standard deviations in the mean, σ/\sqrt{N} . The deviation vector methods are all rescaled. The constrained ellipsoid method rescales and reorthogonalizes every time $T = 1$, and uses a value of $\epsilon = 10^{-6}$ for the tangent-vector inference [Eq. (5.59)]. The error goal is a fractional error of 10^{-10} per step.

Method	λ_1	λ_3
unconstrained deviation vector	0.1610 ± 0.00050	
unconstrained Jacobian	0.1608 ± 0.00050	-0.2618 ± 0.00053
constrained deviation vector	0.1608 ± 0.00051	
constrained Jac. (1 tangent vector)	0.1605 ± 0.00048	
restricted Jacobian	0.1607 ± 0.00048	-0.2614 ± 0.00055
constrained ellipsoid	0.1605 ± 0.00050	-0.2617 ± 0.00051

appears in Table 5.1, which gives error estimates based on integrations using fixed initial conditions and random initial deviations. This table was produced by using an initial point produced from the *final* values of a previous long integration, which avoids any transient effects due to starting at a point not on the attractor. The estimates for the exponents use a final time of $t_f = 10^4$, with 100 randomly chosen values for the deviation vector or initial ball. All the methods agree on the mean exponents to within one standard deviation of the mean. (Recall that we omit the zero exponent associated with the time “degree of freedom.”)

5.4.1 Speed

The various methods for calculating the exponents differ significantly in their execution time, as shown in Table 5.2. Generally speaking, the deviation methods are faster than their Jacobian method counterparts, which is no surprise—the deviation vector methods involve fewer differential equations. More surprising is the performance penalty for the restricted Jacobian method. This is the result of a significantly smaller typical step-size in the adaptive integrator needed to achieve a particular error tolerance. The restricted Jacobian may result in a system of equations that is more difficult to integrate because of the elimination of simple degrees of freedom with the potentially complicated solutions to the constraint derivative equations $\partial C_k / \partial y_i = 0$ [Eq. (5.57)]. On the other hand, the performance penalty of the restricted Jacobian method is probably worth the gain in robustness, as discussed below. Moreover, for other systems (e.g., the system shown in Figs. 5.14 and 5.15), the restricted Jacobian method is comparable in speed to the other Jacobian methods.

Table 5.2: Timing comparison for different Lyapunov exponent methods applied to the forced damped pendulum. The times (on a 2 GHz Pentium 4) for a final time of $t_f = 10^4$ are in seconds: t_1 for the positive exponent λ_1 and t_{1-3} for the negative exponent λ_2 ; we omit the zero exponent (λ_2) for brevity. (We write 1–3 to emphasize that calculating λ_3 also calculates λ_1 as a side-effect.) We consider both the unconstrained [Eq. (5.42)] and constrained [Eq. (5.45)] formulations. The integrations use a C++ Bulirsch-Stoer integrator adapted from [12]. The deviation vector methods are rescaled, and the constrained ellipsoid method rescales and reorthogonalizes every time $T = 1$. The error goal is a fractional error of 10^{-10} per step. The relatively small difference between deviation vector and Jacobian methods is the result of the small number of degrees of freedom; for larger systems (with larger Jacobians) the difference can become large (Chapter 3). We note that the restricted Jacobian method is unusually slow for the forced damped pendulum, but this is not generally the case.

Method	t_1	t_{1-3}
unconstrained deviation vector	2.57	
unconstrained Jacobian	3.65	5.16
constrained deviation vector	3.51	
constrained Jacobian (1 tangent vector)	4.05	
restricted Jacobian	35.3	45.0
constrained ellipsoid	4.30	5.88

5.4.2 Robustness

Numerical methods are more useful if they are relatively insensitive to small changes in implementation details, and the Jacobian methods win in this category. When reorthogonalization occurs every time step, *without* rescaling, the plain Jacobian method is virtually bulletproof. The rescaling in this case can even occur only when the tangent vector norms reach very large or small values, say $\|\xi\| \approx 10^{\pm 100}$. This robustness also applies to the restricted Jacobian method, which is considerably less finicky than any other method for constrained systems, and we recommend its implementation if practical.

Jacobian methods that rescale and reorthogonalize every time T are less robust, since *a priori* we have no knowledge of appropriate values for T . Experimentation in this case is required to find good values of T ; for the Lorenz system, $T = 1$ works well, but $T = 5$ leads to inaccurate estimates for the negative exponent, as seen in Fig. 5.21. It is better to err in the direction of small times, since the Gram-Schmidt procedure is quite robust: even when rescaling occurs on very short timescales—so that the longest axis has almost no chance to outgrow the other principal axes—the Gram-Schmidt method still converges to the correct exponents (Fig. 5.20). Using the Gram-Schmidt algorithm to find the principal axes benefits from a strong feedback mechanism, insuring rapid convergence to the correct axes. Using a very small value for T greatly increases the execution time, of course. A useful prescription in practice is to do a short integration with T chosen to be small compared to any characteristic timescales in the problem, in order to obtain a first estimate for the exponents. We may then choose T to be as large as we like, consistent with the avoidance of unacceptable roundoff

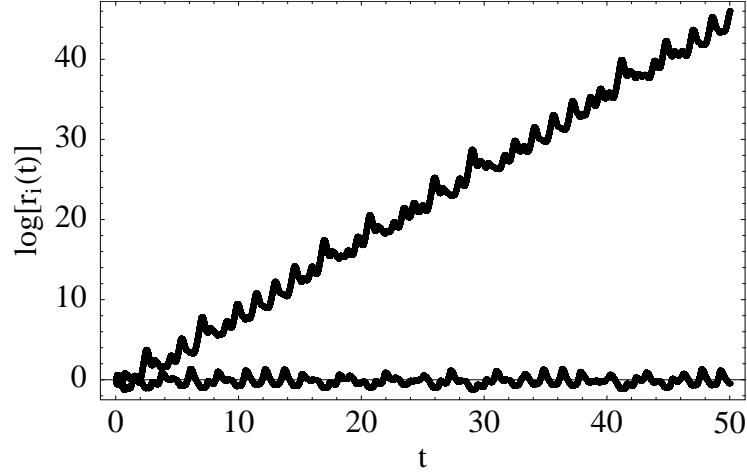


Figure 5.20: The natural logarithms of the two larger ellipsoid axes for the Lorenz system using the Gram-Schmidt algorithm, with the axes rescaled every $T = 10^{-3}$. The largest and smallest directions differ by less than 2% when rescaling this frequently, but the axes nevertheless converge rapidly to the correct directions (as determined by the Jacobian method, Fig. 5.6). Numerical investigations confirm that this robustness persists at least down to $T = 10^{-5}$.

error.

The constrained ellipsoid method is dependent on frequent rescaling to keep the size of the tangent vectors small, since the inference scheme represented by Eq. (5.59) fails for large vector norms. As a result, this method suffers from the complexity of all time T methods, i.e., it requires care in choosing an appropriate value of T . In addition, the value of ϵ in Eq. (5.59) must be chosen carefully to achieve accurate tangent-vector inferences: the method relies on small values of ϵ for accuracy, but values that are too small suffer from roundoff errors. It is advisable to calibrate the value of ϵ so that the largest Lyapunov exponent agrees with the result of a second method (such as the single tangent-vector method or the deviation vector method), as discussed in Chapter 2. Such a calibration was required to produce the values in Table 5.1; the largest exponent calculated using the constrained ellipsoid method differs from the other methods by several standard deviations when using $\epsilon = 10^{-5}$ for the inference, but agrees well when using $\epsilon = 10^{-6}$.

Finally, the deviation vector methods are all very fast, but they are sensitive to the size ϵ_0 of the initial deviation vector. The rescaled methods are particularly inaccurate if the value of ϵ_0 is too small, which leads to roundoff error in the initial size of the deviation vector and can give inaccurate results, as shown in Fig. 5.22. These methods should be used with care, and should always be double-checked with a Jacobian method if possible.

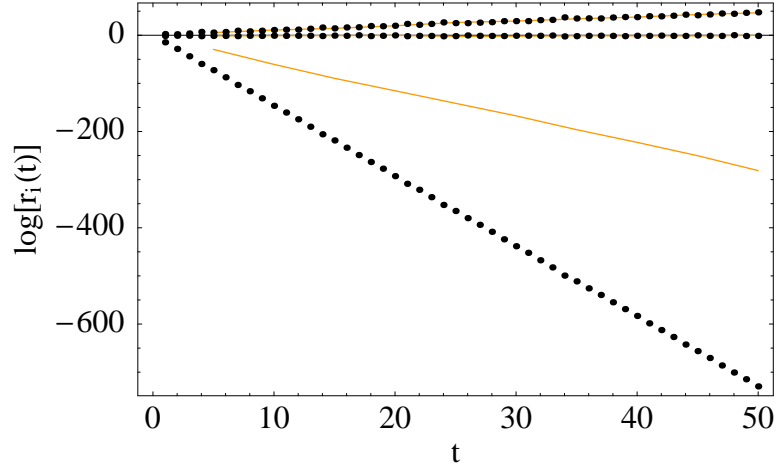


Figure 5.21: The natural logarithms of the ellipsoid axes for the Lorenz system, with reorthogonalization/rescaling every time $T = 1$ (dark dots) and $T = 5$ (light lines). The two larger exponents agree exactly, but the negative exponent is incorrect due to roundoff error, since the smallest axis shrinks from unity to a size of $e^{-5 \times 14.57} \approx 2 \times 10^{-32}$ in a time $T = 5$.

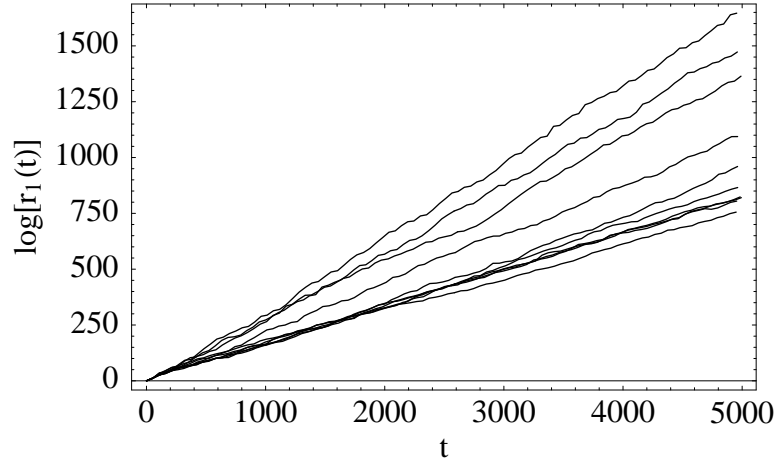


Figure 5.22: The natural logarithms of the largest ellipsoid axis for the constrained forced damped pendulum, calculated using the rescaled deviation vector method for varying sizes of the initial deviation. We vary the size of the initial deviation vector from $\epsilon_0 = 10^{-4}$ (bottom) to $\epsilon_0 = 10^{-13}$ (top). Values of ϵ_0 between 10^{-4} and 10^{-8} agree closely, but smaller values lead to erroneously high values for the Lyapunov exponent. It is important to calibrate the deviation vector method using the Jacobian method (Sec. 5.2.2) if possible.

5.5 Summary and conclusion

Chaotic solutions exist for an enormous variety of nonlinear dynamical systems. Lyapunov exponents provide an important quantitative measure of this chaos. We have presented a variety of different methods for calculating these exponents numerically, both for constrained and unconstrained systems. Both types of systems can be investigated using deviation vector methods or Jacobian methods. Deviation vector methods use the equations of motion to evolve two nearby trajectories in phase space to determine the time-evolution of the small deviation vector joining the trajectories. This family of methods is computationally fast, but yields only the largest exponents, and also suffers from sensitivity to the size of the initial deviations. The Jacobian methods share the use of the Jacobian matrix of the system as a rigorous measure of the local phase-space behavior. They are computationally robust in general, and can be used to determine multiple exponents, but this comes at the cost of execution speed.

Calculating Lyapunov exponents for constrained systems presents a variety of complications, all revolving around the notion of constraint-satisfying deviations: “nearby” trajectories must be chosen carefully to insure that they satisfy the constraints. We have presented several methods for dealing with these complications, including a deviation vector method and two Jacobian methods: the restricted Jacobian method, which eliminates spurious degrees of freedom in the Jacobian by differentiating the constraints; and the constrained ellipsoid method, which uses the full Jacobian matrix to evolve constraint-satisfying tangent vectors. These methods allow the determination of all d Lyapunov exponents for systems with d degrees of freedom.

Acknowledgements

Thanks to Sterl Phinney for encouragement and valuable comments. This work was supported in part by NASA grant NAG5-10707.

Appendix: Ellipsoid axes and the singular value decomposition

In this appendix, we discuss an alternative method for calculating the ellipsoid axes used in the Jacobian method, namely, calculating the ellipsoid axes exactly. The method described seems superior on paper to the Gram-Schmidt technique described in Sec. 5.2.2, but suffers from subtle complications that make it fragile in practice. Nevertheless, within a narrow range of validity (specified below), calculating exact ellipsoid axes provides valuable corroboration of the principal Jacobian method discussed above.

Recall Theorem 1 from Sec. 5.2.2.3, which relates the eigensystem of the matrix $A^T A$ to the ellipsoid spanned by the columns of A . In order to find the axes of an evolving ellipsoid, we could apply Theorem 1 directly, but there is a mathematically equivalent prescription that is numerically virtually bulletproof, namely, the famous *singular value decomposition*:

Theorem 2 *Let A be a nonsingular $n \times n$ matrix. Then there exist orthonormal $n \times n$ matrices U and V , and a diagonal matrix S , such that*

$$A = USV^T. \quad (5.62)$$

This is the singular value decomposition (SVD) of A , and the values s_i in $S = \text{diag}(s_1, \dots, s_n)$ are the singular values.

Since V is an orthogonal matrix, we have $V^T = V^{-1}$, so that Eq. (5.62) is equivalent to $AV = US$. Geometrically, this means that the image of the unit ball V is equal to an ellipsoid whose i th principal axis is given by s_i times the i th column of U . V in this context is a special ball, but the image of any unit ball is the same unique ellipsoid. This leads to the following theorem:

Theorem 3 *Let A be a nonsingular $n \times n$ matrix, and let U and S be the matrices resulting from the singular value decomposition of A [Eq. (5.62)]. Then the columns of A span an ellipsoid whose i th principal axis is $s_i \mathbf{u}_i$, where $S = \text{diag}(s_1, \dots, s_n)$ and $\{\mathbf{u}_i\}_{i=1}^n$ are the columns of U .*

We thus see that the singular value decomposition is equivalent to finding the eigensystem of $A^T A$. (See Appendix A in [9] for proofs of these theorems.)

Substituting the singular value decomposition for the Gram-Schmidt procedure leads to a replacement of step (2) from Sec. 5.2.2:

- (2') At various times t_j , replace \mathbf{U} with the orthogonal axes of the ellipsoid defined by \mathbf{U} , using the singular value decomposition. This can be done either every time T , for some suitable choice of T , or every time the integrator takes a step. *It is essential to order the principal axes consistently.* We recommend sorting the axes so that $s_1 \geq s_2 \geq \dots \geq s_n$.

Unfortunately, this prescription behaves badly when rescaling is necessary, as shown in Fig. 5.23. The underlying cause of this is a fundamental property of the singular value decomposition: it is only unique up to a permutation of the ellipsoid axes. If we adopt an ordering based on the axis lengths, we can refer, for example, to the longest axis as axis 1. During any particular time period, axis 1 may grow or shrink; the only requirement is that it be the fast-growing axis on average. Unfortunately, rescaling the axes causes this ordering method to fail: if axis 1 should happen to contract between rescaling times, then the ordering based on length leads to incorrect axis labels, since axis 1 is no longer the longest axis. Even worse, when ordering by axis length, the length of the longest axis

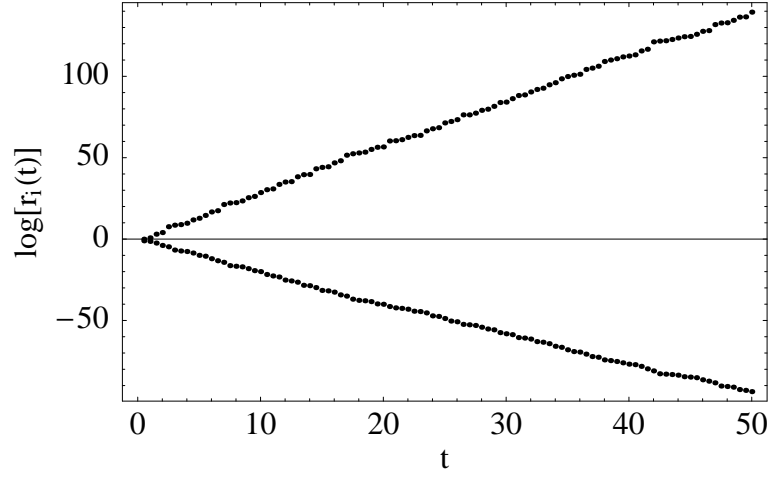


Figure 5.23: The natural logarithms of the two larger ellipsoid axes for the Lorenz system using the singular value decomposition. The axes are rescaled every $T = 0.5$ to exaggerate the deviations from the correct results, but any rescaling causes the SVD method to fail (see text). Compare to unrescaled SVD (Fig. 5.24) and the Gram-Schmidt method with frequent rescaling (Fig. 5.20).

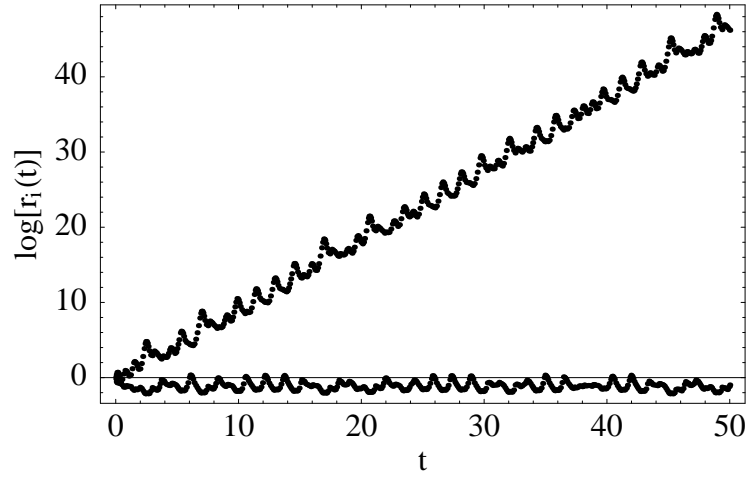


Figure 5.24: The natural logarithms of the two larger ellipsoid axes for the Lorenz system using the singular value decomposition, without rescaling. The results agree well with the Gram-Schmidt method (Fig. 5.6). Rescaling (which is always necessary if we approach the floating point limits of $\sim 10^{\pm 308}$) ruins the agreement.

is always added to the running sum for the largest Lyapunov exponent, while the length of the smallest axis always contributes to the smallest exponent. This selection bias leads to systematic errors, guaranteeing overestimates for the absolute values of both the exponents (Fig. 5.23).

If the system is not rescaled, there is still some initial ambiguity in axis labels, but once axis 1 has grown sufficiently large it is very unlikely ever to become smaller than the other axes. Thus, after an initial expansion and contraction phase that establishes the ordering, the axis labels remain fixed, and the results of the (unrescaled) SVD method agree well with Gram-Schmidt (Fig. 5.24).

It should be possible in principle to follow the axis evolution by tracking the continuous deformation of the ellipsoid. This would mean assigning labels to the axes and then ensuring, e.g., that axis 1 at a later time is indeed the image of the original axis 1. This method would require following the system over very short timescales to guarantee the correct tracking of axes, and even then is likely to be fragile and error-prone. Because of these complications, we recommend the simpler Gram-Schmidt process, which has proven to be reliable and robust in practice.

Bibliography

- [1] E. Lorenz, J. Atmospheric Science **20**, 130 (1963).
- [2] M. Colonna and A. Bonasera, Phys. Rev. E **60**, 444 (1999).
- [3] Y. Y. Yamaguchi and T. Iwai, Phys. Rev. E **64**, 066206 (2001).
- [4] G. A. Gottwald and I. Melbourne, `nlin.CD/0208033`.
- [5] V. B. Ryabov, Phys. Rev. E **66**, 016214 (2002).
- [6] R. O. Vallejos and C. Anteneodo, Phys. Rev. E **66**, 021110 (2002).
- [7] J. D. Barrow and J. Levin, `nlin.CD/0303070`.
- [8] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, England, 1993).
- [9] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems* (Springer, New York, 1997).
- [10] M. D. Hartl, <http://www.michaelhartl.com/software/>.
- [11] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
- [12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, England, 1992).
- [13] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 21 (1980).
- [14] T. Damour, Phys. Rev. D **64**, 124013 (2001).
- [15] T. Damour, P. Jaranowski, and G. Schäfer, Phys. Rev. D **62**, 084011 (2000).
- [16] T. Damour and G. Schäfer, Nuov. Cimento **101**, 127 (1988).
- [17] R. Genzel, C. Pichon, A. Eckart, O. E. Gerhard, and T. Ott, Mon. Not. Royal Astron. Soc. **317**, 348 (2000).

Afterword

We collect here some lessons learned while working on this thesis, some opinions, and selected advice on a variety of topics—including a numerical cautionary tale. We address especially the role played by computation in the numerical study of the systems we consider. Given the informality of this section, and in a break from the style used elsewhere in this thesis, we (that is, I) switch here to the first person singular.

Lessons and advice

- Don't use C for scientific computing.

Although good for speed, C is seriously sub-optimal for scientific computing, with a dearth of mathematical functions and a terribly long development/debug cycle. I wasted countless hours tracking down subtle heisenbugs¹ that usually resulted from some seemingly innocuous change clobbering a block of memory somewhere in the bowels of the program.

- C++ is decent for scientific computing.

C++ addresses some of C's most serious deficiencies (although it becomes a terribly bloated language in the process). The Standard Template Library (STL) container classes are especially great, effectively replacing the inflexible C array data type and alleviating many of the memory-management hassles of C. Still, having to write `pow(x, 1.5)` to raise x to the $3/2$ power (as in C) is insulting.² Expressiveness matters, and C++ doesn't have it for mathematics. I'd seriously consider using Fortran 90/95 for the numerical code (but not for anything else) if I had it to do over again.

- Learn to use a graphical debugger.

Despite the pain of debugging C and C++, it is infinitely worse without DDD, the Data Display Debugger, at your side. By the way, when compiling a C++ program with the GNU C++ compiler, use `g++ -ggdb foo.cpp` in place of `g++ -g foo.cpp`. This allows the debugger to inspect the contents of STL vectors as if they were ordinary C arrays.

¹**heisenbug** /hi:'zen-buhg/ n. A bug that disappears or alters its behavior when one attempts to probe or isolate it. (From the *Jargon File*, <http://catb.org/jargon/html/entry/heisenbug.html>.)

²For a laugh, try writing `pow(x, 3/2)` instead.

- Scripting languages are good (even when they're bad).

Everyone should learn a scripting language. Since they are interpreted (or at most compiled to bytecode), they are slower than C/C++ for numerical tasks, but for most anything else their large libraries and short development/debug cycle are a godsend. The results in Chapter 3 would never have happened without the Perl programming language. As wonderful as it is, Perl is actually terrible. I have since converted to Python, which is much better than Perl, and should be learned by all.

- *Mathematica* is worth it.

Expensive, occasionally buggy, and so very proprietary—*Mathematica* has some significant drawbacks. Unfortunately, it is so powerful that it's worth putting up with the problems. The implementation of the PN equations in *Mathematica* took only a few days, with dozens of extraordinarily useful auxiliary functions written along the way. The various parameterization schemes for the Papapetrou equations, stupidly coded directly in C at great personal cost, were reimplemented inside *Mathematica* in under two days. There's no substitute for having robust equation solvers, root finders, numerical integrators, and symbolic differentiators all readily available. Though difficult after-the-fact debugging is a major flaw in *Mathematica*, the intuitive notebook interface makes for an incredibly fast development/debug cycle. Add to this a flexible and highly expressive programming language, and *Mathematica* becomes a tool too compelling to ignore.

- Use the right language for the job.

I'm definitely still in the learning stage, but already this philosophy has paid major dividends. Using different computer languages for different tasks—even with the added complexity of getting the components to communicate—is a huge win.

The PN integrator is a case in point. The equations of motion, the rendering of the orbits, and various analysis tools (e.g., given a solution to the equations, return the empirical pericenter) are written in *Mathematica*. For many purposes, this is adequate, but calculating Lyapunov exponents requires access to the guts of the integrator, which *Mathematica*'s ODE solver `NDSolve` doesn't allow, and would be hideously slow if implemented in *Mathematica* by hand. These routines are written in C++, as part of a general platform for calculating Lyapunov exponents (as described in Chapter 5), and need only be supplied with the equations of motion, their derivatives, and the Jacobian matrix. The expressions for these for the PN equations are impossibly long, but luckily *Mathematica* can export its native expressions as C code, using the function `CForm`. The resulting code needs to be cleaned up a bit (I/O and string manipulation are not *Mathematica*'s strengths), so I use a series of Python scripts to take the raw *Mathematica* output and convert it to a C++ file, which is then compiled and linked with

the integrator.

The great part is that, by using Unix pipes, temporary files,³ and *Mathematica* options, it is possible to switch between the native *Mathematica* integrator and the C++ back-end on the fly. The speed boost from the C++ routines was high enough that I ended up using them to calculate ordinary orbits, even though they were only strictly necessary for calculating Lyapunov exponents.

- Don't be afraid to learn a new language.

I learned a web scripting language (PHP) for a personal project, and then realized that it wasn't that hard (and was lots of fun) to learn a new language. This led to Perl, Python, C++, and a deeper understanding of *Mathematica*. The productivity boost was more than worth the time spent. Even more important, perhaps, is the *motivation* boost—it's just much more fun to do a project when the tool is right for the job.

- Parameterization matters.

Once under the mistaken impression that setting up the equations of motion was the hard part, I only slowly realized that much of the work comes in implementing the parameterization methods—and a good method makes a *big* difference. The Papapetrou project sputtered along until I implemented the parameterization of initial conditions in terms of the orbit's geometric properties (e, r_p, ι), thereby gaining the ability to (among other things) find bound orbits easily. I shudder to imagine undertaking the broad survey of parameter space in Chapter 3 without this method.

A cautionary tale

And finally, a cautionary tale.

In the early days of the Papapetrou project, I had successfully (I thought) found the Lyapunov exponents for a particular orbit using the Jacobian method, and my initial results were (it seemed) nearly ready to be published, but I thought it prudent to check the Jacobian method using the deviation vector method. I was confident that the Jacobian method worked correctly, since the Lyapunov exponents

1. were “reasonable,” in order of magnitude, based on the results from [1];
2. came symmetrically in pairs $\pm\lambda$, a strong indication that the numerical results correctly captured this key property of Hamiltonian systems (Chapter 2);
3. went to zero as the spin parameter S approached zero, just as they should.

³The ability to create temp files with names guaranteed to be unique is included in virtually all modern computer languages. This is “well-known,” but was not known to this author until recently.

Upon implementing the deviation vector method I discovered that the resulting exponent, when compared with the largest exponent from the Jacobian method,

1. agreed!

By now, convinced that everything was working, I started to move on to other things, but then I realized that I should check the deviation vector method with different values of ϵ (the initial size of the deviation). To my surprise, larger values of ϵ gave different results—and the chaos was gone. Dismayed, I continued testing the deviation vector code, and found that the Lyapunov exponent depended on ϵ : the system was seemingly chaotic for very small ϵ , but the chaos disappeared when it was larger. Moreover, the larger values of ϵ , in the range 10^{-3} – 10^{-7} agreed with each other (and gave vanishing exponents), whereas values in the range 10^{-8} – 10^{-15} gave different results for each ϵ . Evidently, very small values of ϵ were leading to roundoff errors in the deviation vector method, which is essentially a method for taking numerical derivatives—notorious for numerical inaccuracy [2] (an illustration of this phenomenon appears in Chapter 5). I was forced to consider the possibility that the Jacobian method might be wrong, and that I had, purely by chance, initially chosen a (too-small) deviation vector size ϵ that *just happened* to give the same spurious exponent as the (incorrect) Jacobian method. I set out to find an error in the Jacobian matrix.

It is important to realize that the Jacobian matrix for the Papapetrou equations was hand-coded, so debugging it was an arduous process, virtually impossible by visual inspection. I needed some sort of automated test. After considerable experimentation, I finally hit upon the following method (which, with hindsight, is obvious): the error in a finite difference approach should scale as ϵ^2 :

$$\Delta = \mathbf{f}(\mathbf{y} + \delta\mathbf{y}) - \mathbf{f}(\mathbf{y}) - \mathbf{Df} \cdot \delta\mathbf{y} = \mathcal{O}(\epsilon^2), \quad (5.63)$$

where $\epsilon = \|\delta\mathbf{y}\|$. [The variable \mathbf{f} represents the right-hand side of the equations of motion, Eq. (1.37), which I was confident were correct, having tested them extensively by verifying the preservation of the constraints and conserved quantities, and checking several limiting cases.] By calculating Δ for various ϵ , and searching for an ϵ^2 scaling, I would have a sensitive test for a correct Jacobian. I implemented the test. The Jacobian failed it; the error scaled as ϵ . Thus began the search for the error. After many, many⁴ hours, the answer struck: there was a single missing $\mathcal{O}(S^2)$ term in the Jacobian (as noted in the appendix to Chapter 2). Once added, the diagnostic worked, and there was much rejoicing in the land.⁵

Some lessons from this tale:

⁴*many*

⁵The rejoicing did not last long. I quickly realized that I had not addressed the issue of constraints, an issue resolved only after much panic and hand-wringing. That is a story for Chapter 2, with Chapter 5 bringing the method to full fruition. Also, now that my example of chaos was gone, I had to find a new one. The (e, r_p, ι) parameterization method developed in Chapter 3 eventually saved the day, allowing a systematic search of low-pericenter orbits, where chaos (at least for the unphysical $S = 1$ case) is widespread.

- In some ways, numerical analysis shares more in common with experiment than with theory. Teasing a result out of a computer suffers from some of the same fragility and vicissitudes one might encounter with a delicate piece of experimental apparatus. Budget extra time for any computer project, always taking into account Hofstadter's Law.⁶
- Calculate things in more than one way if at all possible.
- Any long, hand-coded equations will have mistakes. A diagnostic, if possible, is invaluable.
- Avoid long, hand-coded equations. Get *Mathematica* to do it. (Then check it with the diagnostic anyway.)

⁶**Hofstadter's Law:** Any computer project will take longer than you think, even when you take into account Hofstadter's Law.

Bibliography

- [1] S. Suzuki and K. Maeda, Phys. Rev. D **55**, 4848 (1997).
- [2] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, England, 1992).